



## Evolutionary history of *Hemerocallis* in Japan inferred from chloroplast and nuclear phylogenies and levels of interspecific gene flow

Hirota, Shun K ; Yasumoto, Akiko A ; Nitta, Kozue ; Tagane, Misa ; Miki, Nozomu ; Suyama, Yoshihisa ; Yahara, Tetsukazu

**Abstract:** The perennial herb genus *Hemerocallis* (Asphodelaceae) shows four flowering types: diurnal half-day, diurnal one-day, nocturnal half-day, and nocturnal one-day flowering. These flowering types are corresponding to their main pollinators, and probably act as a primary mechanism of reproductive isolation. To examine how the four flowering types diverged, we reconstructed the phylogeny of the Japanese species of *Hemerocallis* using 1615 loci of nuclear genome-wide SNPs and 2078 bp sequences of four cpDNA regions. We also examined interspecific gene flows among taxa by an Isolation-with-Migration model and a population structure analysis. Our study revealed an inconsistency between chloroplast and nuclear genome phylogenies, which may have resulted from chloroplast capture. Each of the following five clusters is monophyletic and clearly separated on the nuclear genome-wide phylogenetic tree: (I) two nocturnal flowering species with lemon-yellow flowers, *H. citrina* (half-day flowering) and *H. lilioasphodelus* (one-day flowering); (II) a diurnal one-day flowering species with yellow-orange flowers, *H. middendorffii*; (III) a variety of a diurnal half-day flowering species with reddish orange flowers, *H. fulva* var. *disticha*; (IV) another variety of a diurnal half-day flowering species with reddish orange flowers, *H. fulva* var. *aurantiaca*, and a diurnal one-day flowering species with yellow-orange flowers, *H. major*; (V) a diurnal half-day flowering species with yellow-orange flowers, *H. hakuunensis*. The five clusters are consistent with traditional phenotype-based taxonomy (cluster I, cluster II, and clusters III-V correspond to *Hemerocallis* sect. *Hemerocallis*, *Capitatae*, and *Fulvae*, respectively). These findings could indicate that three flowering types (nocturnal flowering, diurnal one-day flowering, and diurnal half-day flowering) diverged in early evolutionary stages of *Hemerocallis* and subsequently a change from diurnal half-day flowering to diurnal one-day flowering occurred in a lineage of *H. major*. While genetic differentiation among the five clusters was well maintained, significant gene flow was detected between most pairs of taxa, suggesting that repeated hybridization played a role in the evolution of those taxa.

DOI: <https://doi.org/10.1016/j.ympev.2021.107264>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-205409>

Journal Article

Published Version



The following work is licensed under a Creative Commons: Attribution 4.0 International (CC BY 4.0) License.

Originally published at:

Hirota, Shun K; Yasumoto, Akiko A; Nitta, Kozue; Tagane, Misa; Miki, Nozomu; Suyama, Yoshihisa; Yahara, Tetsukazu (2021). Evolutionary history of Hemerocallis in Japan inferred from chloroplast and nuclear phylogenies and levels of interspecific gene flow. *Molecular Phylogenetics and Evolution*, 164:107264.

DOI: <https://doi.org/10.1016/j.ympev.2021.107264>



Contents lists available at ScienceDirect

## Molecular Phylogenetics and Evolution

journal homepage: [www.elsevier.com/locate/ympev](http://www.elsevier.com/locate/ympev)Evolutionary history of *Hemerocallis* in Japan inferred from chloroplast and nuclear phylogenies and levels of interspecific gene flowShun K. Hirota<sup>a,\*</sup>, Akiko A. Yasumoto<sup>b,1</sup>, Kozue Nitta<sup>c</sup>, Misa Tagane<sup>d</sup>, Nozomu Miki<sup>e</sup>, Yoshihisa Suyama<sup>a</sup>, Tetsukazu Yahara<sup>d</sup><sup>a</sup> Field Science Center, Graduate School of Agricultural Science, Tohoku University, 232-3 Aza-yomogida, Naruko Onsen, Osaki, Miyagi 989-6711, Japan<sup>b</sup> Department of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland<sup>c</sup> Department of Environmental Science, School of Life and Environmental Science, Azabu University, 1-17-71 Fuchinobe, Chuo, Sagami-hara, Kanagawa 252-5201, Japan<sup>d</sup> Department of Biology, Faculty of Science, Kyushu University, 744 Motoooka, Fukuoka 819-0395, Japan<sup>e</sup> Graduate School of Systems Life Sciences, Kyushu University, 744 Motoooka, Fukuoka 819-0395, Japan

## ARTICLE INFO

## Keywords:

cpDNA capture  
Hybridization  
Flower color  
Flowering time  
MIG-seq  
SNPs

## ABSTRACT

The perennial herb genus *Hemerocallis* (Asphodelaceae) shows four flowering types: diurnal half-day, diurnal one-day, nocturnal half-day, and nocturnal one-day flowering. These flowering types are corresponding to their main pollinators, and probably act as a primary mechanism of reproductive isolation. To examine how the four flowering types diverged, we reconstructed the phylogeny of the Japanese species of *Hemerocallis* using 1615 loci of nuclear genome-wide SNPs and 2078 bp sequences of four cpDNA regions. We also examined interspecific gene flows among taxa by an Isolation-with-Migration model and a population structure analysis. Our study revealed an inconsistency between chloroplast and nuclear genome phylogenies, which may have resulted from chloroplast capture. Each of the following five clusters is monophyletic and clearly separated on the nuclear genome-wide phylogenetic tree: (I) two nocturnal flowering species with lemon-yellow flowers, *H. citrina* (half-day flowering) and *H. lilioasphodelus* (one-day flowering); (II) a diurnal one-day flowering species with yellow-orange flowers, *H. middendorffii*; (III) a variety of a diurnal half-day flowering species with reddish orange flowers, *H. fulva* var. *disticha*; (IV) another variety of a diurnal half-day flowering species with reddish orange flowers, *H. fulva* var. *aurantiaca*, and a diurnal one-day flowering species with yellow-orange flowers, *H. major*; (V) a diurnal half-day flowering species with yellow-orange flowers, *H. hakuensis*. The five clusters are consistent with traditional phenotype-based taxonomy (cluster I, cluster II, and clusters III-V correspond to *Hemerocallis* sect. *Hemerocallis*, *Capitatae*, and *Fulvae*, respectively). These findings could indicate that three flowering types (nocturnal flowering, diurnal one-day flowering, and diurnal half-day flowering) diverged in early evolutionary stages of *Hemerocallis* and subsequently a change from diurnal half-day flowering to diurnal one-day flowering occurred in a lineage of *H. major*. While genetic differentiation among the five clusters was well maintained, significant gene flow was detected between most pairs of taxa, suggesting that repeated hybridization played a role in the evolution of those taxa.

## 1. Introduction

Since the study of Darwin (1862), a wealth of evidence has been accumulated that adaptation to different pollinator groups promotes the divergence of floral traits and leads to speciation (reviewed in Schiestl and Johnson, 2013). In the process of adaptive divergence and speciation mediated by pollinators, floral displays attracting animal

pollinators co-vary with the preferences and perceptual abilities of pollinators (Dyer et al., 2012). Under this co-variation, a change in a floral display trait can cause pollinator shifts and promote premating reproductive isolation between nascent species (Bradshaw and Schemske, 2003). Floral anthocyanin is a well-studied example of such a trait and repeated anthocyanin gain/loss, which changed flower color drastically, promoted pollinator-mediated divergence in *Mimulus*

\* Corresponding author.

E-mail address: [shun.hirota@gmail.com](mailto:shun.hirota@gmail.com) (S.K. Hirota).<sup>1</sup> Shared first author.<https://doi.org/10.1016/j.ympev.2021.107264>

Received 18 December 2020; Received in revised form 22 June 2021; Accepted 12 July 2021

Available online 14 July 2021

1055-7903/© 2021 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

(Cooley et al., 2011), *Petunia* (Gübitz et al., 2009), *Aquilegia* (Hodges and Derieg, 2009), and *Ipomoea* (Streisfeld and Rausher, 2009). Flowering time can also facilitate rapid speciation by affecting both fitness and assortative mating pleiotropically as a *magic trait* (Matsumoto et al., 2015). Contrastively, postmating reproductive isolation is sometimes weakly developed even between sister species adapted to distinct pollinators, and interspecific hybridization has been found in contact zones, called hybrid zones (Hodges and Arnold, 1994; Wolf et al., 2001). Interspecific hybridization often leads to adaptive introgression (Suarez-Gonzalez et al., 2018) and contributes to new combinations of floral traits that can promote hybrid speciation (Rieseberg, 1997; Rieseberg et al., 1995). Such introgression between sister species adapted to distinct pollinators is a complicated process. A better understanding of this process requires studies that combine, highly-resolved phylogenies, detailed observations on pollinator visits, and reliable estimation of gene flow between the species (van der Niet and Johnson, 2012).

The genus *Hemerocallis* L. (Hemerocallidoideae, Asphodelaceae) is a suitable model to study pollinator-mediated speciation (Rodríguez-Enriquez and Grant-Downton, 2013). *Hemerocallis* is a group of perennial herbs, consisting of about 14 species diverged in East Asia (Chen and Noguchi, 2000; Hotta, 2016; Hu, 1969; Hwang and Kim, 2012). In Japan, six species are distributed (Hotta, 2016). *Hemerocallis* shows four flowering types: diurnal half-day, diurnal one-day, nocturnal half-day, and nocturnal one-day flowering (Chen and Noguchi, 2000; Matsuoka and Hotta, 1966). These flowering types correspond to flower color and their main pollinators. Diurnal half-day flowering species, *H. fulva* (L.) L. and *H. hakuunensis* Nakai, start to flower early in the morning and close flowers one or two hours after sunset. While *Hemerocallis fulva* has reddish orange flowers and is mainly pollinated by swallowtail butterflies (Hirota et al., 2012), *H. hakuunensis* has yellow-orange flowers, and is pollinated by bees (Kang and Chung, 1997a). Diurnal one-day flowering species, *H. middendorffii* Trautv. & C.A.Mey. and *H. major* (Baker) M.Hotta, start to flower early in the morning and close flowers early in the next morning. Both *Hemerocallis middendorffii* and *H. major* have yellow-orange flowers. *Hemerocallis middendorffii* is mainly pollinated by bumblebees (Lelej and Kupianskaya, 2000). The nocturnal half-day flowering species *H. citrina* Baroni and the nocturnal one-day flowering species *H. lilioasphodelus* L. start to flower at dusk and close flowers early in the next morning (*H. citrina*) or in the next evening (*H. lilioasphodelus*). Both *H. citrina* and *H. lilioasphodelus* have lemon-yellow flowers with sweet scent that is lacking in diurnally flowering species. *Hemerocallis citrina* is mainly pollinated by nocturnal hawkmoths (Hirota et al., 2012). In contrast to the strong differentiation among species in floral traits, postmating reproductive isolation is weakly developed in *Hemerocallis*. Different species of *Hemerocallis* can be easily crossed by hand-pollination and F1 hybrids are highly fertile (Kawano and Noguchi, 1973; Matsuoka and Hotta, 1966; Yasumoto and Yahara, 2006, 2008). Also, phenotypic intergradation has been recorded in natural hybrid zones of some pairs of species (*H. fulva* var. *aurantiaca* and *H. citrina* var. *vespertina*, Hasegawa et al., 2006; *H. middendorffii* and *H. yezoensis* (= *H. lilioasphodelus* var. *lilioasphodelus*), Kawano, 1961; *H. fulva* var. *disticha* and *H. citrina* var. *vespertina*, Kawano and Noguchi, 1973). Thus, premating isolation by flowering time and pollinator preferences may function as a primary mechanism of reproductive isolation in *Hemerocallis*.

While *Hemerocallis* has been studied as a model of pollinator-mediated speciation (Hirota et al., 2012, 2013, 2019), phylogenetic relationships among species of *Hemerocallis* still remain unclear. Based on cpDNA phylogenetic analysis, Noguchi and De-yuan (2004) claimed that the nocturnal half-day flowering species, *H. citrina*, derived three times independently from nocturnal one-day flowering ancestors and Noguchi et al. (2004) suggested that *H. citrina*, *H. flava* (= *H. lilioasphodelus*), and *H. middendorffii* were either polyphyletic or paraphyletic. In both studies, however, only a few cpDNA markers were used and the bootstrap values for the cpDNA phylogenetic trees were low. On the other hand, a recently constructed nuclear genome-based

phylogeny of *Hemerocallis* revealed that samples of *H. citrina* and *H. lilioasphodelus* and samples of *H. middendorffii* form two different monophyletic clusters (Murakami et al., 2020). It is well-known that a maternally inherited chloroplast gene-based phylogeny does not always correspond to a nuclear gene-based phylogeny due to chloroplast capture (Rieseberg and Soltis, 1991) or incomplete lineage sorting (Takahashi et al., 2001). The former occurs under a directional gene flow of cpDNA, and the latter is a stochastic process which potentially occurs in a group that evolved through rapid adaptive radiation (Fior et al., 2013). Thus, the inconsistency between the chloroplast phylogeny (Noguchi and De-yuan, 2004; Noguchi et al., 2004) and the nuclear phylogeny (Murakami et al., 2020) may be either of these cases. For phylogenetic reconstruction, Noguchi and De-yuan (2004) and Noguchi et al. (2004) used chloroplast markers only but Murakami et al. (2020) used genome-wide SNPs only so that it still remains uncertain whether their different conclusions are due to inconsistencies of chloroplast and nuclear phylogenies or due to the use of different sample sets. To better understand the evolutionary history of *Hemerocallis* and to resolve conflicting hypotheses about phylogenetic tree topology (monophyletic vs. polyphyletic/paraphyletic), it is necessary to analyze chloroplast and nuclear genomes using the same dataset.

In addition to phylogenetic information, analyses of population structure and interspecific gene flow are required to understand the influence of hybridization on the evolutionary history of *Hemerocallis*. Substantial evidence shows phenotypic intergradation in mixed natural populations of *Hemerocallis* (Hasegawa et al., 2006; Kawano, 1961; Kawano and Noguchi, 1973), but no previous study has been conducted on the genetic population structure and interspecific gene flow in *Hemerocallis*.

Here, we examine the evolutionary history of all diploid Japanese species of *Hemerocallis* by phylogenetic and population genetic approaches. First, we reconstruct two phylogenetic trees separately using chloroplast DNA sequences and nuclear genome-wide SNPs. Second, we estimate genetic diversity indices and the population structure based on nuclear genome-wide SNPs. Third, we estimate the degree and direction of interspecific gene flow using nuclear genome-wide SNPs. Based on those data, we answer to the following questions: (1) Does the chloroplast phylogeny correspond to the nuclear genome phylogeny? (2) If there is any inconsistency between chloroplast and nuclear genome phylogenies, which of chloroplast capture and incomplete lineage sorting is more likely to cause the inconsistency? (3) Is there any significant gene flow between taxa of different flowering types? (4) What is the most plausible process for the adaptive divergence among the species of *Hemerocallis* with different flowering types?

## 2. Materials and methods

### 2.1. Plant materials

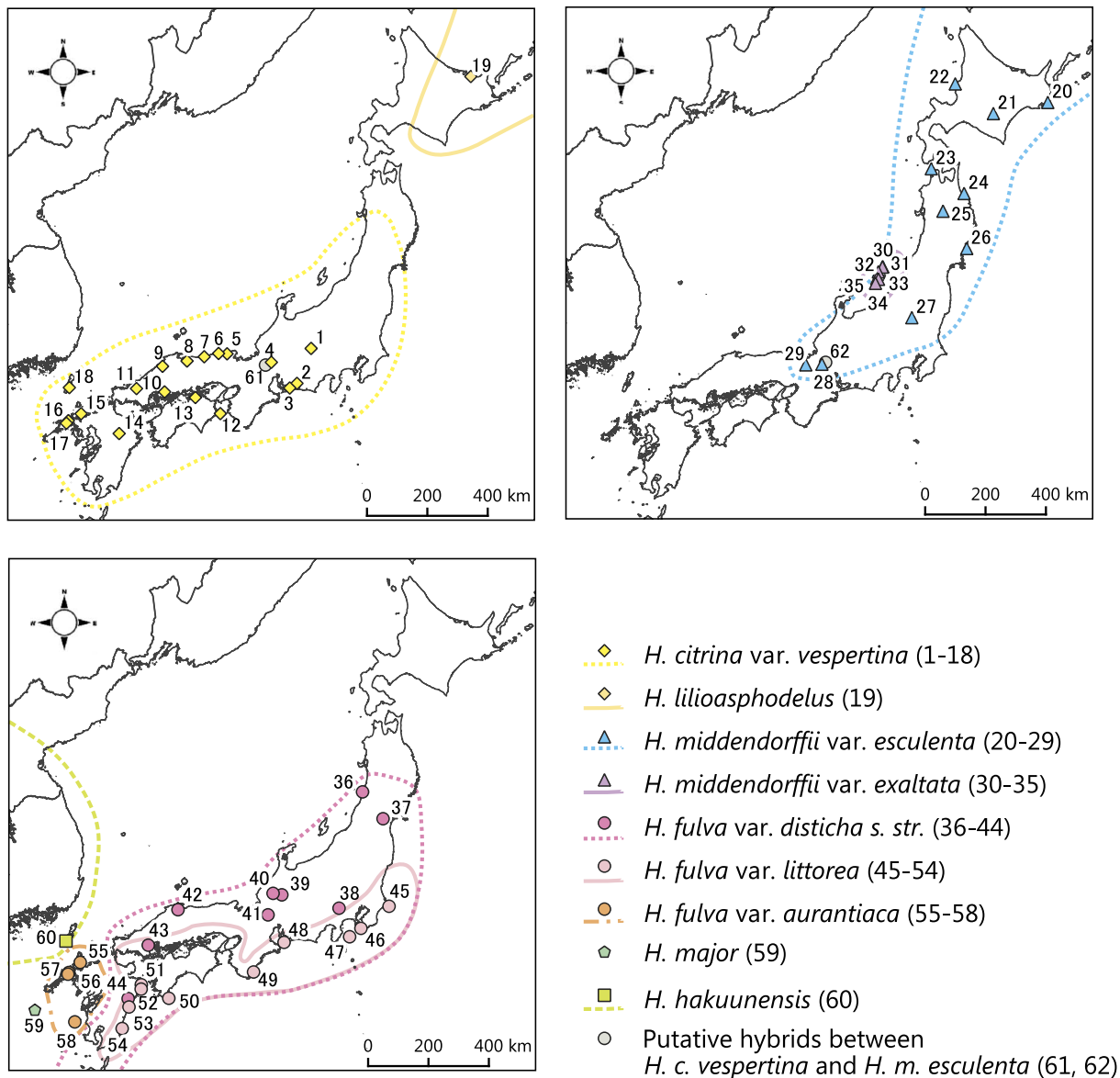
The genus *Hemerocallis* (Asphodelaceae) is divided into three sections: *Hemerocallis*, *Fulvae* Nakai, and *Capitatae* Nakai, distinguished by flowering time, flower color, scents and other morphological characteristics (Hotta, 2016). *Hemerocallis* sect. *Hemerocallis*, including *H. citrina* and *H. lilioasphodelus*, is a group of nocturnally flowering species characterized by lemon-yellow flower color and sweet scent. Both species are probably pollinated by nocturnal hawkmoths (*H. citrina*: Hirota et al., 2012). *Hemerocallis* sect. *Capitatae* includes *H. middendorffii*, a diurnal one-day flowering species, which has yellow-orange flowers with faint scent and is mainly pollinated by bumblebees (Lelej and Kupianskaya, 2000). *Hemerocallis* sect. *Fulvae* includes the diurnal half-day flowering species *H. fulva*, which has scentless reddish orange flowers mainly pollinated by swallowtail butterflies (Hirota et al., 2012). *Hemerocallis major* and *H. hakuunensis* are placed in *H. sect. Fulvae* (Hotta, 2016) or sect. *Hemerocallis* (Murakami et al., 2020). *Hemerocallis major* shows diurnal one-day flowering and *H. hakuunensis* shows diurnal half-day flowering; both have yellow-orange flowers.

*Hemerocallis hakuunensis* is mainly pollinated by bees (Kang and Chung, 1997a). The pollinators of *H. major* have not been reported. On Danjo Islands, which is the only natural habitat of *H. major*, diurnal and nocturnal moths were abundant but large butterflies and bees were rare (Kato et al., 1967; Miyata, 1973; Yamaguchi et al., 1973). Thus, *H. major* may be pollinated by diurnal and nocturnal moths.

## 2.2. Sampling and DNA extraction

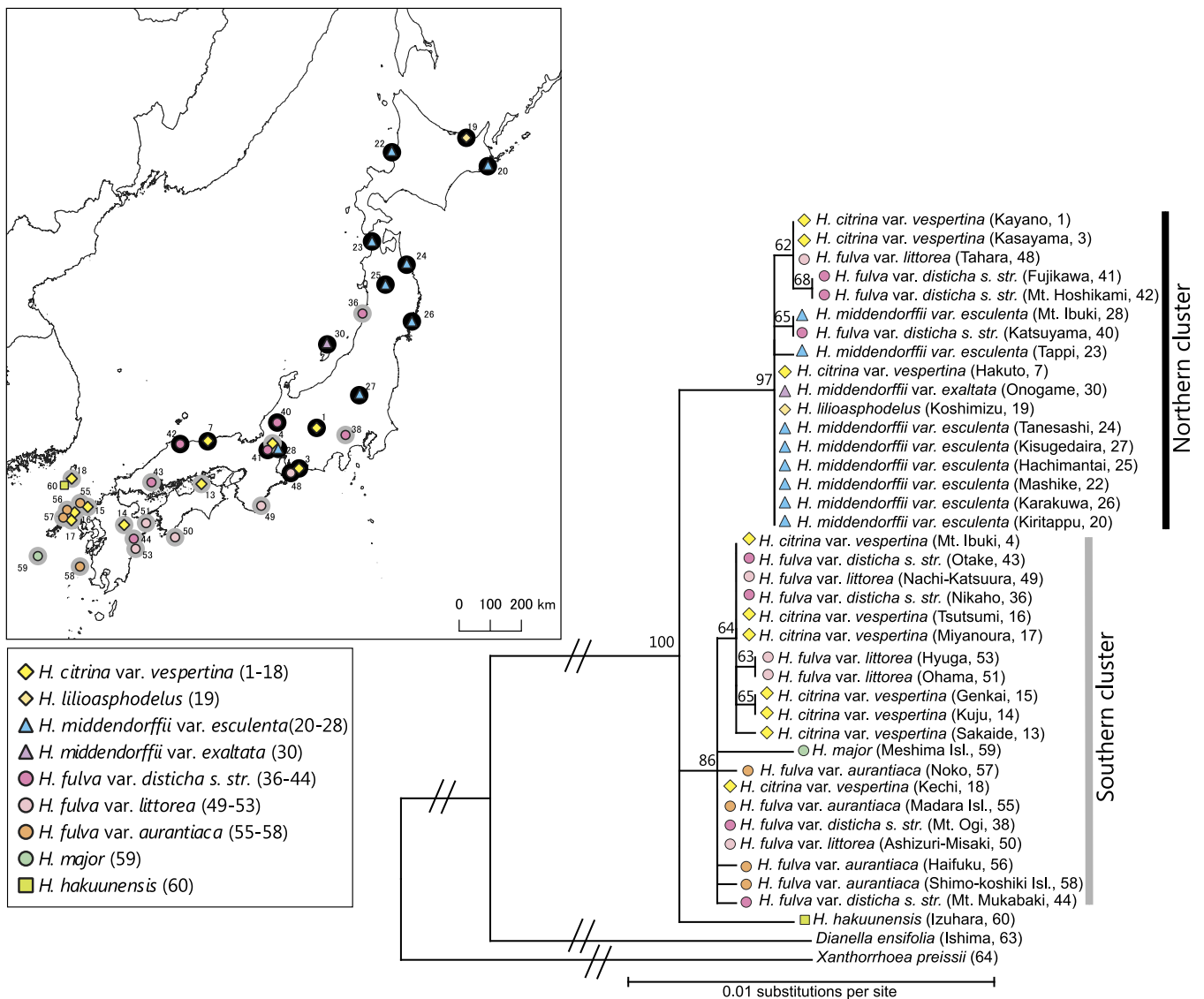
Leaves of the *Hemerocallis* species were sampled from 2001 to 2016 in Japan (Fig. 1, Table S1). In total, 238 individuals were collected from 60 natural populations. These include eight taxa of *Hemerocallis* including six species, two additional varieties and two groups of putative hybrids between *H. citrina* var. *vespertina* (H.Hara) M.Hotta and *H. middendorffii* var. *esculenta* (Koidz.) Ohwi from Mt. Ibuki (Table S1). While Hotta (2016) and Matsuoka and Hotta (1966) distinguished an inland form and a littoral form of *Hemerocallis fulva* as var. *disticha*

(Donn ex Sweet) M.Hotta and var. *littorea* (Makino) M.Hotta, the two varieties are morphologically indistinguishable. Although we at first distinguished var. *disticha* and var. *littorea* by the difference of habitats, as a conclusion of this study, var. *littorea* is treated as a synonym of var. *disticha*. Thus, hereafter, when we mention var. *disticha*, it includes also var. *littorea*. However, in some of the Results sections, particularly figures and tables, we refer to the inland and littoral form of *H. fulva* var. *disticha* as var. *disticha* s. str. and var. *littorea* for convenience. On Mt. Ibuki, *H. citrina* var. *vespertina* is distributed on the mountainside (pop. 4) and *H. middendorffii* var. *esculenta* is found at the summit (pop. 28). In both populations, we found a few individuals whose flowering time, flower color, and/or shape of scape were clearly different from the typical phenotype of each species. In this study, these putative hybrids are designated as hybrid group 61 (found in pop. 4 of *H. citrina* var. *vespertina*) and hybrid group 62 (found in pop. 28 of *H. middendorffii* var. *esculenta*). Additionally, *Dianella ensifolia* (L.) DC. (Hemerocallidoideae, Asphodelaceae) was collected as an outgroup species. Collected leaves



**Fig. 1.** Sampling populations and distribution of *Hemerocallis* in Japan. The numbers indicate the populations (identified in Table S1). *Hemerocallis citrina*, *H. lilioasphodelus*, *H. middendorffii*, *H. fulva*, and *H. hakuunensis* are distributed in eastern continental Asia as well as on the Japanese archipelago. As a conclusion of this study, *H. fulva* var. *littorea* is treated as a synonym of *H. fulva* var. *disticha*. However, in the figures, we refer to the inland and littoral form of *H. fulva* var. *disticha* as var. *disticha* s. str. and var. *littorea* for convenience. The distribution areas refer to Chung and Kang (1994), Hotta (1986), Hotta (2016) and Matsuoka and Hotta (1966).





**Fig. 2.** Chloroplast phylogeny of *Hemerocallis* and geographic distribution of cpDNA samples. Chloroplast phylogeny was constructed by ML analysis (2078 base pairs, 4 loci and 241 polymorphic sites). Bootstrap support values are shown on the nodes. Bootstrap support values below 50 are not shown. The numbers indicate the populations (Table S1). Small symbols (circles, triangles and squares) and their colors indicate species/varieties and their locations in Japan. The background colors of the small symbols indicate the northern (black) or southern (gray) clusters in the cpDNA tree.

were dried with silica gel and stored at room temperature. Voucher specimens were deposited in the Herbarium of Kyushu University (FU). Total DNA was extracted from fresh or silica gel-dried leaves using either a DNeasy Plant Mini Kit (Qiagen) or the CTAB method (Doyle and Doyle, 1990). Extracted DNA of *Xanthorrhoea preissii* Endl. (Xanthorrhoeaceae, Asphodelaceae) was provided from the Kew DNA Bank (reference No. 36236) as an additional outgroup for the MIG-seq analysis.

### 2.3. PCR and sequencing of chloroplast DNA regions

Sequences of chloroplast DNA (cpDNA) regions were determined for 38 individuals of *Hemerocallis* and one individual of *D. ensifolia*. Corresponding sequences of *X. preissii* were obtained from the chloroplast whole genome sequence at GenBank: accession no. KX822774.1. For DNA extraction, we selected one individual from each of the 38 representative populations (Table S1) because the sequences of cpDNA regions of *Hemerocallis* have low genetic diversity (Noguchi and De-yan, 2004; Noguchi et al., 2004). Four non-coding cpDNA regions, *trnL* intron, *trnL-trnF*, *rbcl-atpB* and *psbA-trnH*, were amplified using primers

listed in Table S2. New primers for *rbcl-atpB* were designed for the outgroup species *D. ensifolia* because the universal primers often amplified a non-target region instead of the *rbcl-atpB* region. The primers were designed based on the sequences of the *rbcl* and *atpB* genes of *Dianella*, *Hemerocallis*, and *Xanthorrhoea* from GenBank. The PCR reaction was mainly performed with TaKaRa Ex Taq (Takara Bio) with the PCR cycle consisting of 2 min at 95 °C, followed by 30 cycles of 1 min at 95 °C, 1 min at 55 °C, 1 min at 72 °C, and a final extension of 7 min at 72 °C. The PCR products were purified with an ExoSAP (USB). Except for *psbA-trnH*, the purified PCR products of each cpDNA region were sequenced using the same primers as in the PCR reaction with a BigDye Terminator v 3.1 Cycle Sequencing Kit and an ABI Prism 3100 automated sequencer (Applied Biosystems). For the *psbA-trnH* region, one informative site was very close to the forward primer *psbA\_1f*, and thus we used the different reverse primer *trnH\_2r*, which is slightly more upstream than *trnH\_1r*. All sequence data were registered to DNA Data Bank of Japan (DDBJ) under accession nos. LC566709–LC566825.

## 2.4. MIG-seq analysis for nuclear genome-wide datasets

Multiplexed ISSR genotyping by sequencing (MIG-seq, [Suyama and Matsuki, 2015](#)) was used for *de novo* SNP detection. The MIG-seq library was prepared following the slightly modified protocol of [Suyama and Matsuki \(2015\)](#). The 1st PCR was conducted using MIG-seq primer set 1. The 1st PCR products were diluted 50 times with deionized water and used for the 2nd PCR. The 2nd PCR was conducted using primer pairs including tail sequences, adapter sequences for Illumina sequencing, and the index sequence of a reverse primer to identify each individual sample. The 2nd PCR products were multiplexed in the size range of 300–800 bp. The sequencing was performed by an Illumina MiSeq platform (Illumina, San Diego, CA, USA) using a MiSeq Reagent Kit v3 (150 cycle, Illumina). According to [Suyama and Matsuki \(2015\)](#), we skipped the sequencing of the first 17 bases of read 1 (anchor region for 2nd PCR primer, ISSR primer, and anchor region in the 1st primers) and three bases of read 2 (anchor region for 2nd PCR primer) using the ‘DarkCycle’ option of the MiSeq system. Both ends of the fragments and index sequences were read by pair-end sequences (reads 1 and 2) and index sequencing; 80, 94, and six bases of sequences were determined as read 1, read 2, and the index read, respectively. All raw MIG-seq data were deposited at the DDBJ Sequence Read Archive (DRA) with accession number DRA010527.

### 2.4.1. Quality control and *de novo* assembly

The first 14 bases (12 bases of the SSR primer region and two bases of the anchor sequences in the 1st primers) of read 2 sequences were trimmed using the program ‘fastx\_trimmer’ in the FASTX-Toolkit 0.0.14 ([http://hannonlab.cshl.edu/fastx\\_toolkit](http://hannonlab.cshl.edu/fastx_toolkit)). Low-quality reads were removed from raw reads by the ‘quality\_filter’ option of FASTX Toolkit 0.0.14 using the settings of  $q = 30$  and  $P = 40$ . Extremely short reads containing adapter sequence were also removed from raw reads using TagDust 1.13 ([Lassmann et al., 2009](#)). Quality filtered 80 bp-nucleotide sequences from both ends of the fragment were used for SNP discovery. A *de novo* assembly was performed by *ustacks* and *cstacks* (STACKS 1.47, [Catchen et al., 2013](#)). ‘Stack’ is a set of identical sequences and several of those stacks were merged to form putative loci. First, using the *ustacks* program, each stack was piled within a sample setting the minimum stack coverage at  $m = 3$ , and a set of putative loci within a sample was formed by comparing the stacks allowing for two mismatches between stacks within a locus ( $M = 2$ ). Second, the *cstacks* program merged the putative loci of each sample allowing for two mismatches ( $n = 2$ ) and produced the catalog of putative loci in the entire samples. The default settings for the other options were used. The interspersed repeats and low complexity DNA sequences contained in the catalog of putative loci were treated as missing data using RepeatMasker 4.0.7 ([Smit et al., 2015](#)). The sequences of the repeat-masked putative loci (hereafter referred to as loci) were used as a reference for short-read mapping. Among the repeat-masked putative loci, 24 loci showed high homology (blastn e-value  $< 1e-4$ ) to the complete chloroplast genome of *H. citrina* ([Ou et al., 2020](#), GeneBank Accession No. MN872235). These loci were removed after the mapping process. For short-read mapping, the quality filtered reads from each individual were separately mapped to the repeat-masked putative loci using BWA-MEM alignment 0.7.17 ([Li and Durbin, 2009](#)) with default settings. The sequence alignment/map format files (SAM/BAM) were processed to identify genotypes using SAMtools 1.7 and Bcftools 1.6 ([Li et al., 2009](#)). Then, the genotyped sites with coverage depth less than 10 reads were filtered out.

### 2.4.2. SNP calling and data preparation

The MIG-seq data was obtained from 238 individuals of *Hemerocallis*. Unfortunately, the outgroup samples, *D. ensifolia* and *X. preissii*, had to be eliminated from the data because they did not contain any shared SNP locus with *Hemerocallis*. Next, two datasets, one composed of sequences and another composed of SNPs, were prepared as follows. For the sequence dataset, we parsed the merged genotype file to provide the

sequence data retaining both biallelic SNP sites and non-variable sites with VCFtools 0.1.15 ([Danecek et al., 2011](#)). Four different filtering criteria were applied for quality control of sequence data. First, any SNP site where one of two alleles had less than three counts was filtered out because it is difficult to distinguish polymorphisms from sequencing errors when the minor allele count of SNPs is too low ([Roesti et al., 2012](#)). Second, loci containing SNPs with high heterozygosity ( $H_o \geq 0.6$ ) were removed because excess of heterozygosity potentially resulted from artifactual loci built from several paralogous genomic regions. Third, after performing a Hardy-Weinberg equilibrium test for each population with eight or more individuals, we excluded loci where allele frequencies deviated from the Hardy-Weinberg equilibrium at  $P < 0.01$  for two or more populations ([Larson et al., 2014](#)). This procedure effectively filtered out the loci of the haploid organelle genomes (i.e. the loci of the mitochondria genome, because the loci of the chloroplast genome were already removed before). Fourth, nucleotide sites with a genotyping rate  $< 80\%$  were eliminated. For the SNP-only dataset, we extracted all SNPs from the sequence dataset, removed loci with minor allele frequency  $< 5\%$ , and randomly selected only one SNP site per locus to avoid the inclusion of linked SNPs.

For calculating genetic diversity indices of the nuclear genome-wide sequence, both of the sequence dataset and the SNP-only dataset were used (see next section). For phylogenetic analysis, two additional sequence datasets (genotyping rate  $\geq 60\%$  and  $\geq 30\%$ ) were used along with the above-described sequence dataset (genotyping rate  $\geq 80\%$ ). This is because the resolution of phylogenetic trees tends to be higher when the tree is constructed by a larger dataset even when it includes loci with lower genotyping rates ([Wagner et al., 2013](#)). Among the three datasets, the largest dataset had a genotyping rate  $\geq 30\%$  because it included more loci than the other datasets ( $\geq 60\%$ ,  $\geq 80\%$ ) (Table S4). For outlier detection and population structure analysis, the SNP-only dataset was used. For estimating divergence time and migration rate, a part of the sequence dataset (genotyping rate  $\geq 80\%$ ) was used.

## 2.5. Genetic diversity

The following genetic diversity indices were estimated as additional indicators of past hybridization events. The nucleotide diversity of cpDNA was calculated for species with two or more samples, using DnaSP 6 ([Rozas et al., 2017](#)). The sequence dataset (genotyping rate  $\geq 80\%$ ) and the SNP-only dataset were used to determine total and SNP-based nucleotide diversities of the nuclear genome-wide sequence, respectively; both were computed for species and for populations, using the ‘site-pi’ function of VCFtools ([Danecek et al., 2011](#)). Observed and expected heterozygosities were calculated for species and for populations, using VCFtools for the SNP-only dataset.

## 2.6. Phylogenetic analysis

Multiple alignments of the four cpDNA regions were performed using the program MUSCLE 3.8.1551 ([Edgar, 2004](#)), and alignment columns containing gaps were trimmed using a heuristic selection method based on similarity statistics of trimAl 1.4.rev15 ([Capella-Gutierrez et al., 2009](#)). We used Kakusan 4.0.2016.11.07 ([Tanabe, 2007](#)) to find suitable nucleotide substitution models and partitioning strategies for the nucleotide datasets. The data was subdivided into four components: *rbcl-atpB* (713 bp), *psbA-trnH* (525 bp), *trnL* intron (481 bp) and *trnL-trnF* intergenic spacer (359 bp), and then run through Kakusan4. The AICc criterion ([Sugiura, 1978](#)) was used to compare Nonpartitioned, Partitioned\_Equal\_Mean\_Rate, and Separate models. The Partitioned\_Equal\_Mean\_Rate (GTR +  $\Gamma$ ), which assumes an equal rate of nucleotide substitution across arbitrarily specified partitions, proved optimal. Maximum likelihood phylogenies were inferred using the software RAXML 8.2.10 ([Stamatakis, 2014](#)), whereby 1000 replicates of parallelized tree search bootstrapping was conducted. For the phylogenetic analyses of the nuclear genome-wide sequences dataset, we also

used a maximum likelihood approach in RAXML 8.2.10. We used a GTR +  $\Gamma$  model of sequence evolution, which requires sequence datasets, and performed 1000 replicates of parallelized tree search bootstrapping.

## 2.7. Outlier detection (SNPs under selection)

To distinguish outlier SNPs that are presumably under natural selection from neutral SNPs in the SNP-only dataset, an outlier test was done with the R package *pcadapt* (Luu et al., 2017). This method using a SNP-based principal component analysis has a low false-positive detection rate and does not need a priori grouping. In this method, a user is required to choose the optimal  $K$  principal components (PCs), typically based on inspection of a scree plot, in which  $K$  is the number of PCs with eigenvalues that depart from a straight line. We chose  $K$  values from 1 to 8 to detect outliers. Any SNP with a  $q$ -value less than 0.01 was considered an outlier. When the SNP-only dataset (67 SNPs) was examined with *pcadapt*, the outlier test with the optimal  $K = 3$  revealed 15 SNPs putatively under divergent selection.

## 2.8. Population structure analysis

The population genetic structure of *Hemerocallis* was examined using individual-based principal component analysis (PCA) using GENODIVE 2.0 (Meirmans and Van Tienderen, 2004) and structure analysis (Pritchard et al., 2000). To estimate genetic variation and structure among taxa, the individual-based PCA was performed for two datasets: 52 neutral SNPs and 15 outlier SNPs. We performed structure analysis using STRUCTURE 2.3.4 (Pritchard et al., 2000) for only the 52 neutral SNPs because the software assumes neutrality. We performed 100 independent runs with a burn-in of 100,000 steps and additional 100,000 steps with the admixture model and estimated log-likelihoods for each number of clusters ( $K = 1$ –20). Optimal  $K$  values were determined by using the Delta  $K$  method of Evanno et al. (2005) in Structure Harvester (Earl and Vonholdt, 2012). Graphical results were obtained using CLUMPAK (Cluster Markov Packager Across K, Kopelman et al., 2015) and R package Pophelper 2.2.9 (Francis, 2017).

## 2.9. Estimation of divergence time and migration rate

Three demographic parameters, population size ( $\theta$  for  $4N_e\mu$ ), migration rate ( $m$ ), and divergence time ( $t$ ) were estimated assuming an Isolation-with-Migration (IM) model using the IMA3 program (Hey, 2010; Hey and Nielsen, 2004). For the IM model, we used the part of the sequence dataset consisting of the haplotypes of the 52 loci with the neutral SNPs, which were detected by the outlier test (average length of the loci:  $84.3 \pm 0.6$  bases, minimum and maximum length are 64 and 92 bases, respectively). We used WhatsHap 0.16 (Martin et al., 2016) to determine the haplotypes of each locus. The haplotype sequences of each locus were aligned by MUSCLE 3.8.1551 (Edgar, 2004), and all columns containing any gap were trimmed using trimAl 1.4.rev15 (Capella-Gutierrez et al., 2009). We collected 100,000 genealogies every 100 MCMC steps, each following a burn-in period of 100,000 steps. The uniform prior limits were set for the models within each species ( $[\theta, m, t] = [10, 50, 1]$ ). The HKY model was applied to all loci, 100 Markov-coupled chains with 0.99 heating curve shape parameter were used. The divergence time was scaled using the geometric mean of the mutation rate per locus per year and assuming a generation time of five years. Because the mutation rate, which is required to estimate the demographic parameters, is not known for *Hemerocallis*, we used nuclear ribosomal internal transcribed spacer (ITS) substitution rates of the herbaceous angiosperms (average:  $5.48 \times 10^{-9}$  substitutions per site per year, the lower bound:  $1.72 \times 10^{-9}$ , the upper bound:  $19 \times 10^{-9}$ , Kay et al., 2006). These substitution rates are consistent with the substitution rates of noncoding spacer/introns of nuclear genes of the *Arabidopsis* species (average:  $4.48$ – $9.86 \times 10^{-9}$ , Huang et al., 2012). In addition, we tested the priors ( $[\theta, m, t] = [10, 50, 25]$ ) with 30 Markov-coupled

chains because Noguchi et al. (2004) suggested that ancestors of *H. middendorffii* might have originated at least 25 million years ago.

Because the number of samples per population was not large enough to perform a population-based analysis, we defined “a population” as one of the following four taxa (a–d) corresponding to the four major clusters of the nuclear genome-wide phylogenetic tree (Fig. 3): (a) *H. citrina* var. *vespertina* (cluster I without *H. lilioasphodelus*), (b) *H. middendorffii* var. *esculenta* (cluster II without *H. middendorffii* var. *exaltata*), (c) *H. fulva* var. *disticha* including var. *littorea* (cluster III), and (d) *H. fulva* var. *aurantiaca* (cluster IV without *H. major*). These four taxa were representative of clusters I to IV, each containing more than 20 samples from multiple localities. Because of the small sample size, the endemic taxa (*H. major* and *H. middendorffii* var. *exaltata*, discussed in 4.4) and the taxa only found at one locality in Japan (*H. lilioasphodelus* and *H. hakuensis* (Cluster V)) were not included. Although *H. middendorffii* var. *exaltata* has 22 samples, we did not include it. Because this endemic variety seems to be monophyletically derived from *H. middendorffii* var. *esculenta* in the nuclear genome-wide phylogenetic tree (Fig. 3), *H. middendorffii* var. *esculenta* could be more suitable as a representative of cluster II. Treating each group as a population, we run models for each combination of the four groups. The ancestry of *Hemerocallis* remained unclear because we could not obtain outgroup information from the MIG-seq analysis, and a topological inconsistency was found between the cpDNA and the nuclear genome-wide phylogenies. Thus, we analyzed each pair of taxa separately. Population structure was previously suggested to have little effect on the estimates (Strasburg and Rieseberg, 2010), although the IM models assume that there is no structure within a population (Hey and Nielsen, 2004).

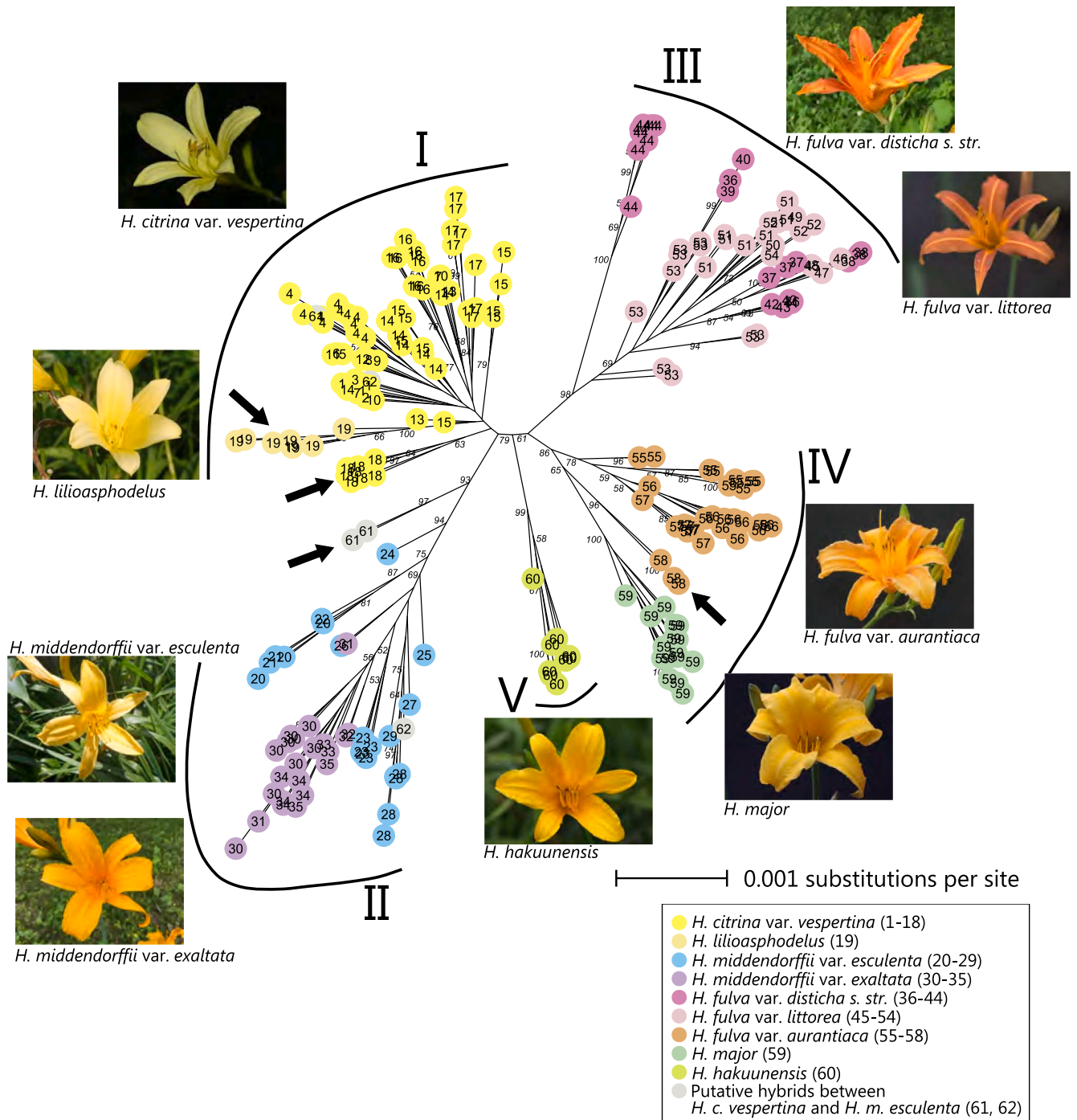
## 3. Results

### 3.1. Genotyping and genetic diversity

The sequences of four cpDNA regions were determined for 40 individuals and, after gaps were trimmed, the total length of the sequences was 2078 bp. We identified 29 and 241 polymorphic sites for *Hemerocallis* and all samples, respectively. There were 16 haplotypes within *Hemerocallis*. Nucleotide diversity of the cpDNA regions across the genus *Hemerocallis* was 0.00263; *H. middendorffii* var. *esculenta* (0.00024) and *H. fulva* var. *aurantiaca* (0.00072) showed nucleotide diversities an order of magnitude smaller than the other taxa (0.00227–0.00254, Table S3).

A total of 68,321,960 raw reads ( $284,674 \pm 7734$  reads per sample) were obtained from MiSeq, and after quality control, 53,974,257 reads ( $224,893 \pm 5295$  reads per sample) remained. The *de novo* assembly produced a total of 82,934 putative loci (6,683,013 bp). Then, 224,376 bp of putative loci were masked by RepeatMasker because of simple repeats (200,608 bp) or low complexity (23,768 bp). Loci obtained from a few individuals were mostly removed during the site filtering process. Consequently, the largest sequence dataset ( $\geq 30\%$  of genotyping rate, 120,579 bp) had 1615 loci shared by at least 72 individuals (Table S4). The smallest dataset ( $\geq 80\%$  of genotyping rate, 14,757 bp) had 200 loci, of which 137 loci (68.5%) contained SNPs, and shared by at least 191 individuals. After filtering out the SNPs which had allele frequencies  $< 5\%$ , 109 SNPs of 67 loci remained. Total nucleotide diversity across *Hemerocallis* was 0.00251. Nucleotide diversity within a taxon was the highest in *H. middendorffii* var. *esculenta* (0.00214) and the lowest in *H. major* (0.00134) (Table S3, but the latter was sampled from only one population). Except for hybrid groups, nucleotide diversity within a population was the highest in the Ohama population of *H. fulva* var. *littorea* and the Izuhara population of *H. hakuensis* (pop. 51, 60, 0.00183, Table S1) and the lowest in the Tahara population of *H. fulva* var. *littorea* (pop. 48, 0.00049, but this nucleotide diversity was calculated from only one sample). As described in Material and Methods 2.2, as a conclusion of this study, *H. fulva* var. *littorea* is treated as a synonym of *H. fulva* var. *disticha*. Hereafter, when we mention var. *disticha*, it includes also var. *littorea*. However, in some of the Results sections,





**Fig. 3.** Nuclear genome-wide phylogeny of *Hemerocallis* based on a ML analysis of the sequences of the MIG-seq data ( $\geq 30\%$  genotyping rate per site, 120,579 base pairs, 1615 loci, and 3457 polymorphic sites). The leaf nodes represent individuals and indicate their population number (Table S1). Arrows indicate populations with a mixture of genetic components from two clusters in the population structure analysis (Fig. 4). Bootstrap support values are shown on the nodes. Bootstrap support values below 50 are not shown.

particularly figures and tables, we refer to the inland and littoral form of *H. fulva* var. *disticha* as var. *disticha* s. str. and var. *littorea* for convenience.

### 3.2. Phylogenetic analysis

The cpDNA phylogenetic tree showed two major geographical groups named northern and southern clusters (Fig. 2). *Hemerocallis middendorffii* var. *esculenta* and var. *exaltata* were included only in the northern cluster, and *H. fulva* var. *aurantiaca* and *H. major* were included

only in the southern cluster. *Hemerocallis hakuunensis* was grouped with neither the northern nor southern cluster. *Hemerocallis citrina* var. *vespertina* and *H. fulva* var. *disticha* were included in both clusters. Most populations of these two taxa belonged to the southern cluster but a few populations (*H. citrina*, pop. 1, 3, 7; *H. fulva* var. *disticha*, pop. 40, 41, 42, 48) belonged to the northern cluster. These seven populations were distributed in the southern area of the northern cluster's distribution range (Fig. 2). Regardless of taxa, the distribution of the northern cluster was biased to the northern and Japan Sea-side areas of Honshu island, and the distribution of the southern cluster was biased to southwestern

Japan (mainly Kyushu and Shikoku islands) except for two population of *H. fulva* var. *disticha* (pop. 36, 38). On Mt. Ibuki, the population of *H. middendorffii* var. *esculenta* (pop. 28) at the higher altitude belonged to the northern cluster, and the population of *H. citrina* var. *vespertina* (pop. 4) at the lower altitude belonged to the southern cluster.

The phylogenetic analysis using the largest dataset of MIG-seq (genotyping rate  $\geq 30\%$ ) produced a topology with the highest bootstrap support and showed five major clusters (Fig. 3). Unlike the cpDNA tree, this MIG-seq tree showed a topology largely consistent with the taxonomy based on floral traits. Cluster I included two nocturnally flowering species, *H. citrina* var. *vespertina* (nocturnal half-day flowering) and *H. lilioasphodelus* (nocturnal one-day flowering), and *H. lilioasphodelus* formed a monophyletic group with 100% of bootstrap value although it was embedded in *H. citrina* (hereafter, for simplification, nocturnal half-day flowering and nocturnal one-day flowering are in some cases collectively referred to as nocturnal flowering). Cluster II included two varieties of a diurnal one-day flowering species, *H. middendorffii* var. *esculenta* and var. *exaltata*. Except for one individual, *H. middendorffii* var. *exaltata* formed a monophyletic group with a bootstrap value of 56%, but *H. middendorffii* var. *esculenta* was not monophyletic if var. *exaltata* was distinguished. The putative hybrids between *H. citrina* var. *vespertina* and *H. middendorffii* var. *esculenta* were placed between clusters I and II: one sample and two samples of putative hybrids from *H. citrina* pop. 4 (hybrid group 61) were assigned to cluster I and a group between cluster I and II, respectively, while one sample of putative hybrids from *H. middendorffii* pop. 28 (hybrid group 62) was assigned to cluster I and the other to cluster II. Cluster III contained a variety of a diurnal half-day flowering species, *H. fulva* var. *disticha*. Cluster IV included another variety of a diurnal half-day flowering species, *H. fulva* var. *aurantiaca*, and the diurnal one-day flowering species *H. major*. The monophyly of *H. major* was supported by a bootstrap value of 100%, and the monophyly of *H. fulva* var. *aurantiaca* except pop. 58 was supported by a bootstrap value of 78%. Pop. 58 of *H. fulva* var. *aurantiaca* was clustered with *H. major* but their monophyly was supported weakly by a bootstrap value of 65%. Cluster V included the diurnal half-day flowering species, *H. hakuunensis*.

Phylogenetic analyses using the smaller datasets (genotyping rate  $\geq 60\%$ ,  $\geq 80\%$ ) revealed major clusters almost corresponding to the major clusters in the phylogenetic tree based on the largest dataset (genotyping rate  $\geq 30\%$ ), but with lower bootstrap supports (Fig. S1). In the phylogenetic analysis using the smallest dataset (genotyping rate  $\geq$

80%), *H. lilioasphodelus* and *H. hakuunensis* were nested within *H. citrina* var. *vespertina*.

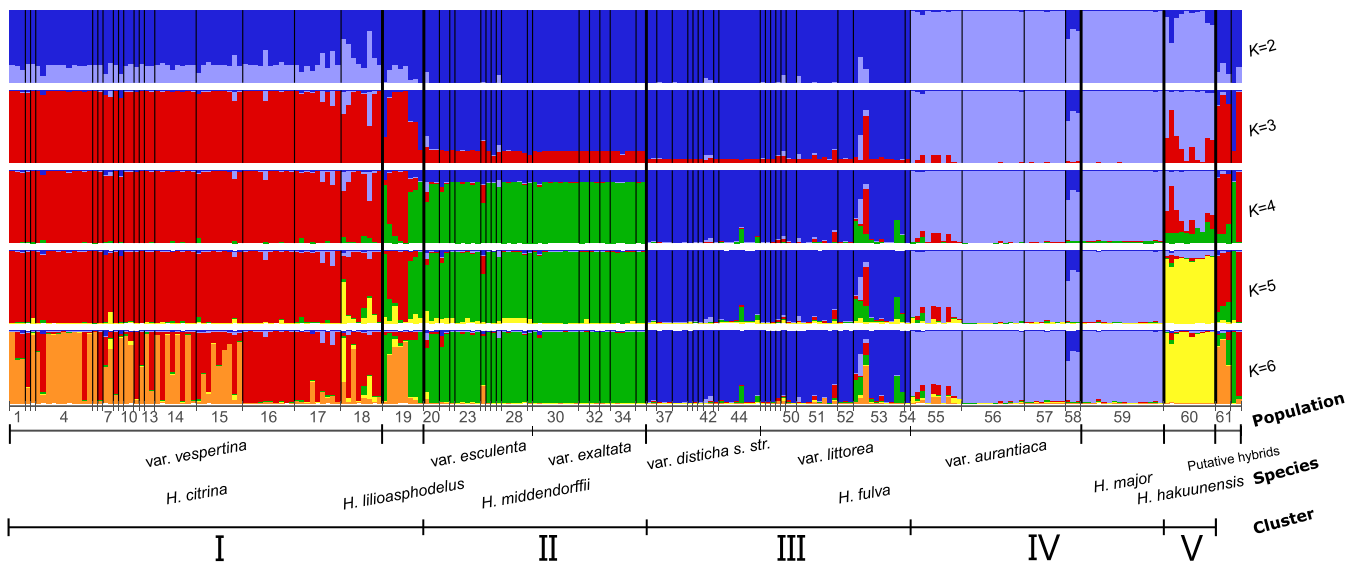
### 3.3. Population structure

The genetic structure examined by STRUCTURE (Pritchard et al., 2000) at  $K = 5$  (Fig. 4) confirmed the five major clusters of the MIG-seq tree (Fig. 3). However, a mixture of two genetic components were found in three populations (pop. 18, 19, 58) and hybrid group 61 (Fig. 4), which were also located at unique positions in the nuclear genome-wide phylogenetic tree (arrows in Fig. 3). First, the population of *H. lilioasphodelus* (pop. 19) had a genetic structure with genetic components of *H. citrina* var. *vespertina* (red in Fig. 4) and *H. middendorffii* (green). Second, the Kechi population of *H. citrina* var. *vespertina* (pop. 18) had genetic components from *H. hakuunensis* (yellow) and *H. citrina* var. *vespertina* (red). Third, the Shimo-koshiki island population of *H. fulva* var. *aurantiaca* (pop. 58) had genetic components from *H. fulva* var. *aurantiaca* (pale blue) and *H. fulva* var. *disticha* (dark blue). Forth, the putative hybrids (61) in the *H. citrina* var. *vespertina* population 4 on Mt. Ibuki had genetic components from *H. citrina* (red) and *H. middendorffii* (green). At  $K = 2$ , the optimal value according to delta  $K$ , the five major clusters of the MIG-seq tree were not separated.

The individual-based PCA using 52 putatively neutral SNPs showed groupings consistent with the five major clusters of the MIG-seq phylogenetic tree using the largest dataset (genotyping rate  $\geq 30\%$ ) although some samples of *H. middendorffii* and *H. fulva* var. *disticha* were overlapped (Fig. S2A). In the PCA diagram using 15 outlier SNPs, however, *H. middendorffii* was well separated from *H. fulva* var. *disticha* (Fig. S2B).

### 3.4. Divergence time and migration between taxa

Considering the result of the MIG-seq phylogenetic tree (Fig. 3), we estimated the divergence time between the following four taxa: (a) *H. citrina* var. *vespertina*, (b) *H. middendorffii* var. *esculenta*, (c) *H. fulva* var. *disticha*, and (d) *H. fulva* var. *aurantiaca*. The estimates of divergence time between taxa were largely overlapping, ranging from 0.39 to 0.53 million years (Table 1, Fig. S3). Subsequently, we estimated migration rates between all pairs of the four taxa. As a result, significant migrations were detected in all pairs of the taxa except the migration from *H. middendorffii* var. *esculenta* to *H. fulva* var. *disticha* (Table 1, Fig. S4).



**Fig. 4.** STRUCTURE plot for  $K = 2-6$ .  $K = 5$  best matches the five major clusters of the phylogenetic analysis using the largest MIG-seq dataset.  $K = 2$  has the largest delta  $K$  for our data. Species and populations are separated by broad and narrow vertical black lines. The population number may not be shown when the sample number from the population is three or less. The clusters shown at the bottom correspond to the nuclear genome-wide phylogenetic tree.

**Table 1**

Demographic parameters estimated by Isolation-with-Migration models for each pair of taxa. High point and 95% highest posterior density intervals in parentheses for divergence time in million years ago ( $t$ ), population size in thousands of taxon 1 ( $\theta_1$ ), taxon 2 ( $\theta_2$ ), ancestral population ( $\theta_A$ ), and migration rate from taxon 1 to 2 ( $m_{1 \rightarrow 2}$ ) and from taxon 2 to 1 ( $m_{2 \rightarrow 1}$ ). Significance of migration was tested by likelihood ratio tests (Nielsen and Wakeley, 2001); \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

Taxon 1	Taxon 2	$t$	$\theta_1$	$\theta_2$	$\theta_A$	$m_{1 \rightarrow 2}$	$m_{2 \rightarrow 1}$
<i>H. citrina</i> var. <i>vespertina</i>	<i>H. middendorffii</i> var. <i>esculenta</i>	0.431 (0.311–0.614)	87.1 (64.2–131.5)	82.0 (55.3–122.6)	15.9 (7.0–40.0)	0.675 (0.275–1.575)***	1.375 (0.525–2.925)***
<i>H. citrina</i> var. <i>vespertina</i>	<i>H. fulva</i> var. <i>disticha</i>	0.527 (0.332–0.787)	105.4 (82.7–132.8)	57.7 (39.9–76.8)	31.6 (10.1–58.9)	0.375 (0.125–1.025)**	2.225 (1.175–3.575)***
<i>H. citrina</i> var. <i>vespertina</i>	<i>H. fulva</i> var. <i>aurantiaca</i>	0.497 (0.130–0.769)	110.4 (84.5–143.8)	27.8 (15.4–43.8)	14.2 (0.0–43.8)	1.175 (0.475–2.175)**	5.525 (3.025–8.725)***
<i>H. middendorffii</i> var. <i>esculenta</i>	<i>H. fulva</i> var. <i>disticha</i>	0.398 (0.244–0.680)	76.7 (47.1–121.9)	48.4 (27.7–76.7)	21.3 (7.1–45.8)	0.275 (0.000–1.575)	1.625 (0.525–3.225)***
<i>H. middendorffii</i> var. <i>esculenta</i>	<i>H. fulva</i> var. <i>aurantiaca</i>	0.423 (0.248–0.717)	76.7 (45.2–120.0)	33.4 (17.7–54.4)	15.1 (0.0–36.1)	0.975 (0.375–2.075)***	0.825 (0.175–2.125)**
<i>H. fulva</i> var. <i>disticha</i>	<i>H. fulva</i> var. <i>aurantiaca</i>	0.392 (0.206–0.729)	53.7 (33.0–79.6)	33.0 (17.5–51.1)	30.4 (9.7–61.5)	2.225 (1.025–3.875)***	1.928 (0.575–3.425)***

Two of the six pairs showed a significant asymmetric migration. The migration rate from *H. fulva* var. *disticha* to *H. citrina* var. *vespertina* was significantly larger than the migration rate of the reverse direction in that the lower bound of the 95% HPD interval of the former ( $m_{2 \rightarrow 1}$ ) exceeded the upper bound of the 95% HPD interval of the latter ( $m_{1 \rightarrow 2}$ ). Similarly, the migration rate from *H. fulva* var. *aurantiaca* to *H. citrina* var. *vespertina* was significantly larger than the reverse direction migration rate. These estimates agreed well with the estimates obtained using the larger divergence time prior ( $[\theta, m, t] = [10, 50, 25]$ ), indicating that the estimates of divergence time were robust (Table S5).

#### 4. Discussion

The nuclear genome-wide phylogeny was not consistent with the chloroplast phylogeny: the former showed five major clusters which agreed with the phenotype-based taxonomy (particularly based on floral traits) but the latter corresponded to the geographic distribution (i.e. a northern and a southern cluster). This discrepancy is probably caused by chloroplast capture, as discussed below. The nuclear genome phylogeny is more reliable than the chloroplast phylogeny because the population structure based on the neutral loci showed that genetic differentiation was well maintained among the five major clusters although significant interspecific gene flow was detected between most pairs of taxa. Adaptation of floral traits to pollinators could be the main factor to maintain the genetic differentiation among the five major clusters. This reliable nuclear genome-wide phylogeny allows consideration of the most plausible speciation process in *Hemerocallis* offering an explanation of how the different flowering types and the three sections of *Hemerocallis* (*Capitatae*, *Fulvae*, *Hemerocallis*) have evolved. Our findings could indicate that the different flowering types have diverged early in the evolutionary history of *Hemerocallis* and have been maintained conservatively.

##### 4.1. Inconsistency between chloroplast and nuclear genome-wide phylogeny

In plants, there are many examples of inconsistencies between the phylogenies from chloroplast and nuclear genes attributed to chloroplast capture (Rieseberg and Soltis, 1991) or incomplete lineage sorting (Takahashi et al., 2001). While both explanations are applicable to the inconsistency in *Hemerocallis*, we suggest below that chloroplast capture is more likely to be the case.

Introgression occurs through hybridization and subsequent backcross, and the specific case of cpDNA introgression is known as chloroplast capture (Rieseberg and Soltis, 1991). Chloroplast capture can occur if the chloroplast and nuclear genomes have different modes of inheritance. In general, chloroplast and nuclear genomes are inherited maternally and biparentally, respectively. *Hemerocallis* probably has this mode of inheritance as well. In *Hemerocallis*, introgression is expected

because many of the species can easily hybridize in natural environments (Hasegawa et al., 2006; Kawano, 1961; Kawano and Noguchi, 1973) and also through hand-pollination (Kawano and Noguchi, 1973; Matsuoka and Hotta, 1966; Yasumoto and Yahara, 2006). The significant interspecific gene flow revealed in this study (Table 1, Fig. S4) confirmed that introgressions of nuclear genes did occur in the evolutionary history of *Hemerocallis*. During the process of recurrent backcrossing, not only nuclear genes but also chloroplast genes would have been exchanged between species. The majority of populations of *H. citrina* and *H. fulva* var. *disticha* had the southern cluster type of cpDNA while a few populations of those two species had the northern cluster type of cpDNA (Fig. 2). All other species had only one type of cpDNA (Fig. 2). As a result, *H. citrina* and *H. fulva* var. *disticha* showed much larger nucleotide diversities than other species (Table S3). Those populations of *H. citrina* and *H. fulva* with the northern cluster type of cpDNA may have originated from a hybridization with a species distributed in northern Japan (e.g. *H. middendorffii*), causing chloroplast capture. Incomplete lineage sorting, on the other hand, is provided by stochastic sorting of ancestral polymorphisms (Takahashi et al., 2001). Our results did not support stochastic sorting of ancestral polymorphisms because the inconsistencies between the chloroplast phylogeny and the nuclear phylogeny were not random but geographically biased.

##### 4.2. Phylogeny and flowering types

Among the five major clusters of the nuclear genome-wide phylogenetic tree, cluster I and II correspond to a nocturnal flowering group (nocturnal one-day and half-day flowering) and a diurnal one-day flowering group, respectively, and clusters III–V correspond to a group of diurnal half-day flowering species except for the diurnal one-day flowering species *H. major*. This finding could indicate that the three flowering types (nocturnal flowering, diurnal one-day flowering, and diurnal half-day flowering) diverged early in the evolutionary history of *Hemerocallis* and have been maintained conservatively and, subsequently, diurnal one-day flowering evolved in *H. major* of cluster IV, and nocturnal one-day flowering and nocturnal half-day flowering differentiated in *H. lilioasphodelus* and *H. citrina* of cluster I, respectively.

The ancestral flower state remained unclear because our nuclear genome-wide phylogenetic tree was not rooted, and a root of the cpDNA tree was trichotomous. To infer the ancestral flower state, we need to consider flower traits of the clade sister to *Hemerocallis*. According to a recent molecular phylogenetic study of Xanthorrhoeaceae, now included in Asphodelaceae (McLay and Bayly, 2016), *Hemerocallis* is sister not to *Dianella*, which we used as an outgroup in our phylogenetic analysis, but to *Simethis* Kunth or *Chamaescilla* Mueller ex Benth. *Simethis* is a monotypic genus distributed in western Europe and northern Africa. *Chamaescilla* is a genus including four Australian species. *Hemerocallis*, *Simethis* and *Chamaescilla* compose the hemerocallid



clade, which is sister to the johnsonioid clade (distributed mainly in Australia) of Hemerocallidoideae (McLay and Bayly, 2016). Regarding flowering time, *Chamaecilla corymbosa* flowers diurnally and probably longer than one day (McLay, Todd G. B., personal communications, March 5, 2020). The flower longevity of *Simethis* is unknown but all taxa of the johnsonioid clade are diurnal one-day flowering (Clifford and Conran, 1998). Regarding pollination, *Simethis* and *Chamaecilla* appear to be pollinated by bees (McLay and Bayly, 2016) and the johnsonioid clade contains mainly buzz-pollinated species (Clifford and Conran, 1998). More generally, it was suggested that buzz-pollination by bees is ancestral in Hemerocallidoideae and butterfly or moth pollination evolved later in *Hemerocallis* (Furness et al., 2013).

According to the knowledge summarized above, the common ancestors of *Hemerocallis* may have been diurnal, one-day or longer than one-day flowering, and their flowers may have been mainly pollinated by bees. Among all *Hemerocallis* species, *H. middendorffii* is most similar to the hypothesized ancestor because it is diurnal one-day flowering and mainly pollinated by bees (Lelej and Kupianskaya, 2000). In contrast, all other species do not satisfy these three conditions simultaneously (see in Materials and Methods 2.1). Below, we thus assume that the flowering type of cluster II, including *H. middendorffii*, is the ancestral flower state of *Hemerocallis*. Under this assumption, Fig. 5A shows the most parsimonious hypothesis for the evolutionary history of *Hemerocallis* suggesting that all clusters (I–V) are monophyletic and that nocturnal half-day flowering and diurnal half-day flowering emerged independently from a diurnal one-day flowering ancestor.

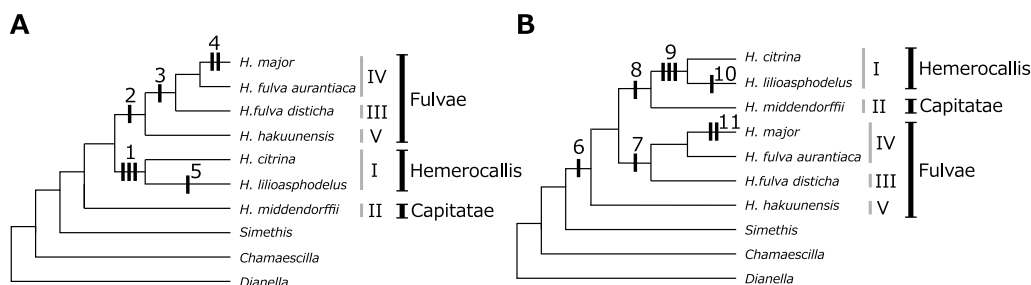
The hypothetical evolutionary history of *Hemerocallis* in Fig. 5 focuses on four flower traits: flower opening time, flower closing time, flower color, and floral scent. Flower opening and closing time together can determine flower duration. Each of these four traits is taxonomically important to classify *Hemerocallis* species (Hotta, 2016) and relevant for the pollination syndrome (van der Pijl, 1961). Moreover, flower opening and closing time of *Hemerocallis* are regulated by different major genes (Nitta et al., 2010). More generally, in plants, flower color and flower scent are often governed by major genes (color: e.g., Bradshaw and Schemske, 2003; Hoballah et al., 2007; scent: e.g., Klahre et al., 2011). Therefore, drastic changes in these four flower traits can be caused by a small number of mutations in major functional or regulatory genes.

Under the most parsimonious hypothesis (Fig. 5A), three of the four traits changed in the common ancestor of cluster I (“1” in Fig. 5A): ancestral diurnal one-day flowering changed to nocturnal half-day flowering (a shift of flower opening time from morning to dusk), yellow-orange flower color changed to lemon-yellow, and scentless flowers changed to flowers with a sweet scent. These changes of flower traits are typically associated with a pollinator shift from diurnal insects to hawkmoths (van der Pijl, 1961). On the sister branch, in the common ancestor of cluster III, IV, and V (“2” in Fig. 5A), diurnal one-day flowering was shortened to diurnal half-day flowering through a shift of flower closing time from the next morning to the evening before. This change may be associated with a specialization to diurnal pollinators. Subsequently, in the common ancestor of cluster III and IV (“3” in

Fig. 5A), flower color changed from yellow-orange to reddish orange. This color change may be associated with a pollinator shift from bees to butterflies because butterflies prefer red flowers to yellowish flowers (Hirota et al., 2012, 2013). Next, within cluster IV (“4” in Fig. 5A), flower color changed from reddish orange to yellow-orange and diurnal half-day flowering was extended to diurnal one-day flowering in the ancestor of *H. major* through a shift of flower closing time from the evening to the next morning. These two changes may be associated with adaptation to a local pollinator environment of Danjo Islands, which is the only natural habitat of *H. major* (discussed in 4.3). Finally, in the ancestor of *H. lilioasphodelus* within cluster I (“5” in Fig. 5A), nocturnal half-day flowering was extended to nocturnal one-day flowering through a shift of flower closing time from the morning to the evening. This could also imply adaptation to a local pollinator environment but our knowledge of pollinators for *H. lilioasphodelus* on the continent (Russia, China and Korea) are too limited to derive any conclusion.

Fig. 5A suggests that two pollinator shifts occurred in *Hemerocallis*: from bees to hawkmoths (“1” in Fig. 5A) and from bees to butterflies (“3” in Fig. 5A), each associated with a change in flower color from yellow-orange to lemon-yellow and to reddish orange, respectively. The former, the evolution of hawkmoth pollination in *Hemerocallis* (“1” in Fig. 5A) is different from the previously known cases of *Aquilegia* and *Petunia* (Klahre et al., 2011; Whittall and Hodges, 2007), in which flower color changed from pink or red to white, resulting in a pollinator shift from bees or hummingbirds to hawkmoths. In *Aquilegia* and *Petunia*, a flower color change directly influenced preference of pollinators. In *Hemerocallis*, however, hawkmoths do not show a consistent preference to lemon-yellow flower color (Hirota et al., 2013) but prefer the high contrast of a UV bullseye pattern as a nectar guide (Hirota et al., 2019). The flower color change from yellow-orange to lemon-yellow may thus be related to a change in the contrast of the UV bullseye pattern. The latter, the evolution of butterfly pollination in *Hemerocallis*, directly corresponds to pollinators’ color preference. Butterflies consistently prefer reddish flower color in *Hemerocallis* (Hirota et al., 2012, 2013). Additionally, bees take longer to find red flowers than flowers with more conspicuous colors (Spaethe et al., 2001). Thus, flower color change from yellow-orange to reddish orange has probably promoted a pollinator shift from bees to butterflies.

In *Hemerocallis*, flower colors are mainly attributed to anthocyanins and carotenoids (Griesbach and Batdorf, 1995). In some cases, drastic changes of flower color caused by major gene mutations result in a pollinator shift (Bradshaw and Schemske, 2003; Hoballah et al., 2007; Owen and Bradshaw, 2011). Such major genes are often related to transcription factors involved in anthocyanin (e.g. reddish, pink color) or carotenoid (e.g. yellow, orange color) pigmentation synthesis (Quattrocchio et al., 1999; Sagawa et al., 2016). The pollinator shifts from bees to hawkmoths (“1” in Fig. 5A) and from bees to butterflies (“3” in Fig. 5A) were associated with a flower color change from yellow-orange to lemon-yellow and to reddish orange, respectively. The former may be related to a reduction or change of floral carotenoid pigmentation, whereas the latter may be related to an increased



**Fig. 5.** Hypothetical pathways of evolutionary events in floral traits. The ancestral state was assumed as (A) diurnal one-day flowering with yellow-orange flowers and (B) diurnal half-day flowering with yellow-orange flowers. The Roman numerals correspond with cluster I–V. The black vertical bars indicate evolutionary events. The Arabic numerals 1–11 refer to evolutionary events mentioned in the discussion.



carotenoid and largely increased anthocyanin pigmentation.

We also discuss another hypothetical evolutionary history of *Hemerocallis* because our dataset is limited to Japanese *Hemerocallis* species, and does not include the four species endemic to China (*H. multiflora* Stout, *H. nana* Forrest & W. W. Smith, *H. forrestii* Diels and *H. plicata* Stapf, [Chen and Noguchi, 2000](#)) and the two species endemic to Korea (*H. hongdoensis* M.G. Chung & S.S. Kang and *H. taeanensis* S.S. Kang & M.G. Chung, [Hwang and Kim, 2012](#)). The four Chinese species were suggested to belong to sect. *Fulvae* ([Matsuoka and Hotta, 1966](#)), and the two Korean species were described under sect. *Fulvae* ([Chung and Kang, 1994](#); [Kang and Chung, 1997b](#)). These species are all diurnal half-day flowering and have uniformly colored orange or yellow-orange flowers, which are common to *H. hakuunensis* of sect. *Fulvae*. Among the six unanalyzed species, the three Chinese species, *H. forrestii*, *H. plicata*, and *H. nana*, are endemic to the highlands of south-western China, especially the northwest of Yunnan ([Chen and Noguchi, 2000](#)), which is known to have high plant diversity containing palaeo-endemisms ([López-Pujol et al., 2011](#)). Thus, *H. forrestii*, *H. plicata*, and *H. nana* may be a group of old origin and could be an ancestral lineage. Considering this possibility, we below reconstruct an evolutionary history of *Hemerocallis* by assuming that *H. hakuunensis*, which has the above character states common to the three Chinese species, is an ancestor.

Under this assumption ([Fig. 5B](#)), flower color changed from yellow-orange to reddish orange in the common ancestor of cluster III and IV (“7” in [Fig. 5B](#)), which may be associated with a pollinator shift from bees to butterflies. On the sister branch, in the common ancestor of cluster I and II (“8” in [Fig. 5B](#)), diurnal half-day flowering was extended to diurnal one-day flowering through a shift of flower closing time from the evening to the next morning. The subsequent evolutionary events in cluster I (“9, 10” in [Fig. 5B](#)) and IV (“11” in [Fig. 5B](#)) occurred in the same order as in [Fig. 5A](#). Interestingly, under this different evolutionary hypothesis ([Fig. 5B](#)), the two pollinator shifts from bees to other insects occurred, just as in [Fig. 5A](#): from bees to hawkmoths (“9” in [Fig. 5B](#)) and from bees to butterflies (“7” in [Fig. 5B](#)). However, the evolutionary path in [Fig. 5B](#) requires an additional evolutionary event compared to the path in [Fig. 5A](#): in [Fig. 5B](#), flowering duration once shortened from diurnal one-day flowering to diurnal half-day flowering when the ancestral state of *Hemerocallis* was derived from closely related genera (*Simethis* and *Chamaesilla*) (“6” in [Fig. 5B](#)) but this event did not occur in [Fig. 5A](#).

Although our MIG-seq tree is unrooted, the estimated divergence time (0.39–0.53 MYA) is consistent with the previous estimation based on cpDNA ([McLay and Bayly, 2016](#)). Therefore, *Hemerocallis* is likely to have experienced rapid adaptive divergence during the Quaternary. [McLay and Bayly \(2016\)](#) showed, using cpDNA regions, that the ancestor of *Hemerocallis* could have diverged from the sister genus of *Simethis* at least 10 MYA. So that, the ancestral state is likely to have been maintained since 10 MYA to 0.39–0.53 MYA when the adaptive radiation happened. During the long-term continuation of the ancestral state, a lot of mutations may have accumulated, resulting in difficulties of reconstructing a rooted tree. While the genus *Hemerocallis* is distributed in East Asia, the sister genus *Simethis* occurs in Western Europe and Northern Africa ([Gianguzzi et al., 2012](#)) and the second closest genus *Chamaesilla* is found in Australia ([McLay and Bayly, 2016](#)). These three geographic ranges are totally separated from each other. Since the history of the separation of these ranges remains uncertain, it is difficult to determine the genus closest to *Hemerocallis* to infer its ancestral state.

To understand and reconstruct the evolutionary history of *Hemerocallis* in greater detail, more information is required: first, a rooted phylogenetic tree; second, more samples of populations and species of *Hemerocallis* from outside of Japan; and third, more information about the flowering types and pollinators of the taxa closely related to *Hemerocallis*.

#### 4.3. Interspecific gene flow

Significant interspecific gene flow was detected between most pairs of taxa. Moreover, biased gene flow was observed from the diurnal half-day flowering species (*H. fulva* var. *aurantiaca* and var. *disticha*) to nocturnal flowering species *H. citrina* ([Table 1](#), [Fig. S4](#)). Biased gene flow was detected only between pairs of taxa whose flowering time overlaps for a few hours in the early evening only ([Hasegawa et al., 2006](#); [Nitta et al., 2010](#)). This biased gene flow probably resulted from a biased pollen flow from early to later flowering species caused by the behavior of crepuscular hawkmoths. Assuming that pollinators have no preference for floral traits, it was theoretically predicted that biased pollen flow from diurnal half-day flowering to nocturnal half-day flowering species always happens because of pollen carry-over ([Matsumoto et al., 2013](#)). Crepuscular hawkmoths, one of the main pollinators of *H. citrina*, foraged rather opportunistically and showed no significant preference to flower color and scent ([Hirota et al., 2013](#)) despite of their preference for a high bullseye contrast pattern, which is stronger in flowers of *H. citrina* than *H. fulva* ([Hirota et al., 2019](#)). This provides a possible explanation for the biased gene flow from *H. fulva* to *H. citrina*. By contrast, unbiased bidirectional gene flow was observed between the other pairs of species whose flowering time overlapped over half a day or longer. The longer overlap period may diminish the effect of biased pollen flow from early to later flowering species.

While significant gene flow was detected between most pairs of taxa, the population structure based on neutral SNPs showed that genetic differentiation was well maintained among the five major clusters ([Fig. 4](#)). In general, under significant gene flow, closely related taxa are expected to show low genetic differentiation ([Cooper et al., 2010](#)). However, when local adaptations by divergent selection have been going on, genetic differentiation among taxa can be maintained not only for a relatively short time (e.g. *Aquilegia*, [Li et al., 2014, 2019](#)) but also for a longer time period (e.g. *Helianthus*, [Strasburg and Rieseberg, 2008](#)). In this study, 15 outlier SNPs were found presumably under natural selections, indicating that divergent selection could have promoted local adaptations which possibly have played a role for maintaining the observed genetic difference even in the face of the significant gene flow detected in this study as well. In *Hemerocallis*, adaptation of floral traits to pollinators could have been more important than adaptation of vegetative traits to abiotic environments because the five clusters are clearly distinguished by floral traits but not by vegetative traits. The impact of interspecific gene flow and local adaptation on the evolutionary history of *Hemerocallis* is, however, to be studied in more detail using larger datasets.

In spite of the well-maintained genetic structure of the five major clusters, at the population level, three cases of introgression between clusters were found with the population structure analysis in populations 18, 19, and 58 ([Fig. 4](#)). First, population 19 of *H. lilioasphodelus* had a genetic structure intermediate between *H. citrina* var. *vespertina* (cluster I) and *H. middendorffii* var. *esculenta* (cluster II) ([Fig. 4](#)). [Kawano \(1961\)](#) reported a large natural hybrid swarm of *H. lilioasphodelus* and *H. middendorffii* at Otanoshike in the eastern part of Hokkaido and suggested that these species are normally isolated by habitats but they can easily hybridize bidirectionally once their populations get into contact with each other because their flowering times totally overlap (both of them are one-day flowering). Second, population 18, the Kechi population of *H. citrina* on Tsushima island, had genetic components not only from *H. citrina* (cluster I) but also from *H. hakuunensis* (cluster V) ([Fig. 4](#)). [Hotta et al. \(1984\)](#) recorded large phenotypic variations of flower morphology and flowering time in the Kechi population of *H. citrina*, suggesting the introgression of *H. hakuunensis* genes into *H. citrina*. The flowering times of these two species overlapped for a few hours in the early evening, like in the case of *H. citrina* and *H. fulva*, so that biased gene flow from *H. hakuunensis* to *H. citrina* may have happened. Third, population 58 of *H. fulva* var. *aurantiaca* on the Shimokoshi island had genetic components from *H. fulva* var. *aurantiaca*

(cluster IV) and from var. *disticha* (cluster III) (Fig. 4). While there have been no previous reports of introgressions between these two varieties, it is likely that gene flow from var. *disticha* to var. *aurantiaca* (pop. 58) happened at some moment.

In *H. citrina* population 4 and *H. middendorffii* population 28 at Mt. Ibuki, we found a few individuals of the putative hybrids (hybrid group 61, 62), but no sign of introgression between pop. 4 and 28 was observed in the structure analysis (Fig. 4). In addition, nucleotide diversities and heterozygosities of the parental populations (pop. 4 and 28) were similar to other populations of *Hemerocallis* (Table S1). These putative hybrids had genetic components of *H. citrina* (cluster I) as well as *H. middendorffii* (cluster II). It is notable that one individual from hybrid group 62 (in pop. 28 of *H. middendorffii*) was almost identical with *H. citrina* (Fig. 4) although it was growing in the *H. middendorffii* population. This individual may be a derivative of backcrossing from *H. citrina* (pollen donor) to a hybrid in pop. 28 of *H. middendorffii* (pollen recipient). Further study on the cpDNA of this individual is needed because it is not determined from our sample.

According to Hotta (1986), there is a hypothesis that *H. major* evolved from a hybridization between *H. fulva* var. *aurantiaca* and *H. citrina*. However, this hypothesis contradicts the result of the population structure analysis (Fig. 4), which showed that the genetic components of *H. major* are almost identical to *H. fulva* var. *aurantiaca* and there was no sign of gene flow from *H. citrina*. Interestingly, there was no individual with an *H. fulva* var. *aurantiaca*-like phenotype (reddish orange flowers and diurnal half-day flowering) in the *H. major* population (yellow-orange flowers and diurnal one-day flowering, Yasumoto, personal observation) although the genetic components of *H. major* are largely identical to *H. fulva* var. *aurantiaca* (Fig. 4). *Hemerocallis major* may have evolved and diverged from *H. fulva* var. *aurantiaca* (Fig. 3) because of the long-time isolation from *H. fulva* and the local pollinator environment on Danjo Islands. In the world, *H. major* is only naturally distributed on Danjo Islands. Danjo Islands have been geologically isolated from Kyushu island and the Continent by the sea and have developed their own unique fauna and flora with endemic species and varieties (Ito et al., 2017; Toriba, 1986; Uematsu et al., 1973; Yamaguchi and Ejima, 1973). According to the records of the fauna on Danjo Islands, diurnal and nocturnal moths are abundant but large butterflies and bees are rare (Kato et al., 1967; Miyata, 1973; Yamaguchi et al., 1973).

#### 4.4. Taxonomic insights

The nuclear genome-wide phylogenetic tree agreed with the phenotype-based taxonomy (Hotta, 2016; Matsuoka and Hotta, 1966), except for the result that *H. fulva* var. *disticha* s. str. and var. *littorea* were not separated clearly (Fig. 3). *Hemerocallis* sect. *Hemerocallis*, *Capitatae*, and *Fulvae* corresponded with cluster I, II, and III-V, respectively (Fig. 5). Our results also largely agreed with another MIG-seq study of *Hemerocallis* by Murakami et al. (2020) which found that the following three clusters were clearly separated: the first cluster composed of *H. middendorffii* varieties (*H. sect. Capitatae*, equivalent to our cluster II), the second cluster composed of *H. fulva* varieties (*H. sect. Fulvae*, equivalent to our cluster III), and the third cluster composed of *H. citrina* varieties, *H. lilioasphodelus* varieties, *H. hakuunensis* and *H. major* (*H. sect. Hemerocallis*, closest to our cluster I). In contrast to Murakami et al. (2020), our nuclear genome-wide phylogenetic tree showed two new additional clusters (IV and V) and *H. hakuunensis* and *H. major* did not belong to *Hemerocallis* sect. *Hemerocallis* (cluster I) but to *H. sect. Fulvae* (cluster III-V) as suggested by Hotta (2016). In our result, *H. hakuunensis* formed the monophyletic cluster V, which was distinctly separated from the other clusters (Fig. 3). Also, *H. major* formed a monophyletic clade which was clearly separated from *H. fulva* var. *aurantiaca* in cluster IV. These differences between the two studies can be attributed to the limited sample size of *H. hakuunensis* (three), *H. major* (one) and *H. fulva* var. *aurantiaca* (none) in Murakami et al. (2020). In contrast, our study

used 10, 16, and 33 samples of *H. hakuunensis*, *H. major*, and *H. fulva* var. *aurantiaca*, respectively, providing more reliable evidence for the phylogenetic positions of *H. hakuunensis* and *H. major*.

Our nuclear genome-wide phylogenetic tree offers three suggestions for the intraspecific and species-level taxonomy of *Hemerocallis* in Japan. First, *Hemerocallis fulva* var. *disticha* s. str. and var. *littorea* are not separated (in cluster III, Fig. 3). Based on this result, we treat var. *littorea* as a synonym of var. *disticha* as follows.

*Hemerocallis fulva* L., Sp. Pl. ed. 2: 462 (1762)

var. *disticha* (Doon ex Sweet) M.Hotta, Acta Phytotax. Geobot. 22: 43 (1966); Fl. Jap. 4b:133 (2016).

var. *littorea* (Makino) M.Hotta, Acta Phytotax. Geobot. 22: 42 (1966), syn. nov.; Fl. Jap. 4b: 133 (2016).

*Hemerocallis fulva* var. *littorea* occurs in coastal habitats and was distinguished by its evergreen leaves from var. *disticha* which is deciduous in winter (Matsuoka and Hotta, 1966). However, the inland strains are often evergreen under cultivation and it is hard to distinguish the two varieties morphologically. Second, *H. fulva* var. *aurantiaca* has highly diverged from *H. fulva* var. *disticha*, and is closer to *H. major*. The distribution of *H. fulva* var. *aurantiaca* and *H. major* is limited to western Kyushu (Hotta, 1986) and both are distinguished by floral color and morphology. Thus, *H. fulva* var. *aurantiaca* may be treated as a distinct species sister to *H. major*. However, it is hard to distinguish *H. fulva* var. *aurantiaca* from *H. fulva* var. *disticha* by floral traits because the two varieties are very similar in having reddish orange tepals with dark tawny mottling and human-unrecognized scent. To describe *H. fulva* var. *aurantiaca* as a species, we need to define additional diagnostic morphological traits. Third, *H. middendorffii* var. *exaltata* endemic to Tobishima and Sado islands was derived from *H. middendorffii* var. *esculenta* widely distributed in the higher elevation of northern Japan. *Hemerocallis middendorffii* var. *exaltata* has a later flowering season and more flowers (10–30 in well-developed plants) per inflorescence than *H. middendorffii* var. *esculenta* (3–10 flowers) (Hotta, 2016; Matsuoka and Hotta, 1966). It is possible that these flowering traits are an adaptation to the local pollinator environment on the islands, like in the case of *H. major* (discussed in 4.3). This result supports the suggestion of Matsuoka and Hotta (1966) that *H. middendorffii* var. *exaltata* is a lineage of *H. middendorffii* adapted to the lowland environment of Tobishima and Sado islands.

## 5. Conclusions

Using 1615 loci of genome-wide SNPs (MIG-seq) and 2078 bp sequences of four cpDNA regions, our study revealed an inconsistency between chloroplast and nuclear genome phylogenies, probably due to chloroplast capture. Significant interspecific gene flow was detected between most pairs of taxa, but the five major clusters are clearly separated on the nuclear genome-wide phylogenetic tree corresponding to floral traits. This could indicate that gene flow among the five clusters is limited at the present time or that pollinator-mediated selection for floral traits is stronger than the effect of gene flow. Among the five major clusters, two clusters corresponded to a nocturnal flowering group and a diurnal one-day flowering group, respectively, and the other three clusters corresponded to a diurnal half-day flowering group except for the diurnal one-day flowering species *H. major*. This finding could indicate that these three flowering groups have diverged early in the evolutionary history of *Hemerocallis*. *Hemerocallis* provides a promising model system to examine the speciation process with gene flow.

## CRedit authorship contribution statement

**Shun K. Hirota:** Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing – original draft, Writing - review & editing, Visualization, Supervision, Funding acquisition. **Akiko A. Yasumoto:** Conceptualization, Methodology, Investigation, Resources, Data curation, Writing - review & editing, Supervision, Funding

acquisition. **Kozue Nitta:** Conceptualization, Methodology, Investigation, Resources, Data curation, Writing - review & editing, Supervision, Funding acquisition. **Misa Tagane:** Investigation, Resources, Data curation, Writing - review & editing. **Nozomu Miki:** Investigation, Writing - review & editing. **Yoshihisa Suyama:** Methodology, Writing - review & editing, Funding acquisition. **Tetsukazu Yahara:** Conceptualization, Methodology, Resources, Writing - review & editing, Supervision, Funding acquisition.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

The authors thank Hirofumi Fujimoto, Taku Fujita, Masahiro Hasegawa, Ryuji Ichihashi, Masashi Igari, Yuji Isagi, Shiro Kobayashi, Masayuki Maki, Tadayoshi Murase, Wataru Ohnishi, Shota Sakaguchi, Masamitsu Tada, Shuichiro Tagane, Hironori Toyama, Gen Yamashita, Kew DNA Bank for their help in sampling materials. We also thank Moeko Hirata, Chika Mitsuyuki, Hidenori Tachida, Koichi Kamiya and Kosuke M. Teshima for their valuable advice on molecular analyses. We also thank Todd G. B. McLay for the information about the floral traits of *Chamaescilla*.

## Funding

The present study was financed by JSPS KAKENHI (Grants-in-Aid for Scientific Research) (grant number: JP17J01165 to SKH, JP24247005 to TY, JP21770025 to AAY, JP23128510 and JP25128711 to KN) and by the Environment Research and Technology Development Fund (JPMEERF20204001 to TY and YS) of the Environmental Restoration and Conservation Agency of Japan.

## Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ympev.2021.107264>.

## References

- Bradshaw, H.D., Schemske, D.W., 2003. Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. *Nature* 426 (6963), 176–178. <https://doi.org/10.1038/nature02106>.
- Capella-Gutierrez, S., Silla-Martinez, J.M., Gabaldon, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25 (15), 1972–1973. <https://doi.org/10.1093/bioinformatics/btp348>.
- Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., Cresko, W.A., 2013. Stacks: an analysis tool set for population genomics. *Mol. Ecol.* 22 (11), 3124–3140. <https://doi.org/10.1111/mec.12354>.
- Chen, X., Noguchi, J., 2000. *Hemerocallis*. *Flora of China* 24, 161–165.
- Chung, M.G., Kang, S.S., 1994. *Hemerocallis hongdoensis* (Liliaceae): A New Species from Korea. *Novon* 4 (2), 94–97. <https://doi.org/10.2307/3391573>.
- Clifford, H.T., Conran, J.G., 1998. Johnsoniaceae. In: Kubitzki, K. (Ed.), *Flowering Plants. Monocotyledons: Liliaceae (except Orchidaceae)*. Springer New York, pp. 336–340.
- Cooley, A., Modliszewski, J., Rommel, M., Willis, J., 2011. Gene duplication in *Mimulus* underlies parallel floral evolution via independent trans-regulatory changes. *Curr. Biol.* 21 (8), 700–704. <https://doi.org/10.1016/j.cub.2011.03.028>.
- Cooper, E.A., Whittall, J.B., Hodges, S.A., Nordborg, M., Joly, S., 2010. Genetic variation at nuclear loci fails to distinguish two morphologically distinct species of *Aquilegia*. *PLoS ONE* 5 (1), e8655. <https://doi.org/10.1371/journal.pone.0008655>.
- Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., Durbin, R., 2011. The variant call format and VCFtools. *Bioinformatics* 27 (15), 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>.
- Darwin, C., 1862. *On the various contrivances by which British and foreign orchids are fertilised by insects*. John Murray, London, UK.
- Doyle, J.J., Doyle, J.L., 1990. Isolation of plant DNA from fresh tissue. *Focus* 12, 39–40.
- Dyer, A.G., Boyd-Gerny, S., McLoughlin, S., Rosa, M.G.P., Simonov, V., Wong, B.B.M., 2012. Parallel evolution of angiosperm colour signals: common evolutionary pressures linked to hymenopteran vision. *Proc. R. Soc. B* 279 (1742), 3606–3615. <https://doi.org/10.1098/rspb.2012.0827>.
- Earl, D.A., vonHoldt, B.M., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4 (2), 359–361. <https://doi.org/10.1007/s12686-011-9548-7>.
- Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32 (5), 1792–1797. <https://doi.org/10.1093/nar/gkh340>.
- Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14 (8), 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>.
- Fior, S., Li, M., Oxelman, B., Viola, R., Hodges, S.A., Ometto, L., Varotto, C., 2013. Spatiotemporal reconstruction of the *Aquilegia* rapid radiation through next-generation sequencing of rapidly evolving cpDNA regions. *New Phytol.* 198, 579–592. <https://doi.org/10.1111/nph.12163>.
- Francis, R.M., 2017. pophelper: an R package and web app to analyse and visualize population structure. *Mol. Ecol. Resour.* 17, 27–32. <https://doi.org/10.1111/1755-0998.12509>.
- Furness, C.A., Conran, J.G., Gregory, T., Rudall, P.J., 2013. The trichotomosulcate asparagoids: pollen morphology of Hemerocallidaceae in relation to systematics and pollination biology. *Aust. Syst. Bot.* 26, 393–407. <https://doi.org/10.1071/sb13046>.
- Gianguzzo, L., Cusimano, D., Bonventre, V., Romano, S., Ilardi, V., 2012. Bio-ecological, phytosociological and conservation aspects of relictual and disjointed populations of *Simethis mattiazzi* (Vandellii) Sacc. (Xanthorrhoeaceae) in the Channel of Sicily. *Acta Bot. Gallica* 159 (3), 303–318. <https://doi.org/10.1080/12538078.2012.737141>.
- Gübitz, T., Hoballah, M.E., Dell'Olivo, A., Kuhlmeier, C., 2009. *Petunia* as a Model System for the Genetics and Evolution of Pollination Syndromes. In: Gerats, T., Strommer, J. (Eds.), *Petunia: Evolutionary, Developmental and Physiological Genetics*. Springer, New York, pp. 29–49.
- Griesbach, R.J., Batdorf, L., 1995. Flower pigments within *Hemerocallis fulva* L. *fm. fulva*, *fm. rosea*, and *fm. disticha*. *HortScience* 30 (2), 353–354.
- Hasegawa, M., Yahara, T., Yasumoto, A., Hotta, M., 2006. Bimodal distribution of flowering time in a natural hybrid population of daylily (*Hemerocallis fulva*) and nightlily (*Hemerocallis citrina*). *J. Plant Res.* 119 (1), 63–68. <https://doi.org/10.1007/s10265-005-0241-3>.
- Hey, J., 2010. Isolation with migration models for more than two populations. *Mol. Biol. Evol.* 27, 905–920. <https://doi.org/10.1093/molbev/msp296>.
- Hey, J., Nielsen, R., 2004. Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* 167, 747–760. <https://doi.org/10.1534/genetics.103.024182>.
- Hirota, S.K., Miki, N., Yasumoto, A.A., Yahara, T., 2019. UV bullseye contrast of *Hemerocallis* flowers attracts hawkmoths but not swallowtail butterflies. *Ecol. Evol.* 9 (1), 52–64. <https://doi.org/10.1002/ecs3.4604>.
- Hirota, S.K., Nitta, K., Kim, Y., Kato, A., Kawakubo, N., Yasumoto, A.A., Yahara, T., Zeil, J., 2012. Relative role of flower color and scent on pollinator attraction: experimental tests using F1 and F2 hybrids of daylily and nightlily. *PLoS ONE* 7 (6), e39010. <https://doi.org/10.1371/journal.pone.0039010>.
- Hirota, S.K., Nitta, K., Suyama, Y., Kawakubo, N., Yasumoto, A.A., Yahara, T., Warrant, E., 2013. Pollinator-mediated selection on flower color, flower scent and flower morphology of *Hemerocallis*: evidence from genotyping individual pollen grains on the stigma. *PLoS ONE* 8 (12), e85601. <https://doi.org/10.1371/journal.pone.0085601>.
- Hoballah, M.E., Gübitz, T., Stuurman, J., Broger, L., Barone, M., Mandel, T., Dell'Olivo, A., Arnold, M., Kuhlmeier, C., 2007. Single gene-mediated shift in pollinator attraction in *Petunia*. *Plant Cell* 19, 779–790. <https://doi.org/10.1105/tpc.106.048694>.
- Hodges, S.A., Arnold, M.L., 1994. Floral and ecological isolation between *Aquilegia formosa* and *Aquilegia pubescens*. *Proc. Natl. Acad. Sci. USA* 91 (7), 2493–2496. <https://doi.org/10.1073/pnas.91.7.2493>.
- Hodges, S.A., Derieg, N.J., 2009. Adaptive radiations: from field to genomic studies. *Proc. Natl. Acad. Sci. USA* 106 (Suppl 1), 9947–9954. <https://doi.org/10.1073/pnas.0901594106>.
- Hotta, M., 1986. *Hemerocallis aurantiaca* group in Northern Kyushu. *Japan. Acta Phytotax. Geobot.* 37, 17–22 (in Japanese).
- Hotta, M., 2016. In: Iwatsuki, K., Boufford, D.E., Ohba, H. (Eds.), *Flora of Japan*, vol. IVb. Angiospermae: Monocotyledoneae (b). Kodansha, Tokyo, pp. 130–134.
- Hotta, M., Ito, M., Okada, I., 1984. Anthesis of the genus *Hemerocallis* and its variation. Special mentions to nocturnal *H. thunbergii* of Tsushima & Hirado islands, western Japan. *Acta Phytotax. Geobot.* 35, 84–83 (in Japanese). <http://doi.org/10.18942/bunruichiri.KJ00001078492>.
- Hu, S.Y., 1969. A new daylily species from Sakhalin Island. *Hemerocallis* J. 23, 42–43.
- Huang, C.-C., Hung, K.-H., Wang, W.-K., Ho, C.-W., Huang, C.-L., Hsu, T.-W., Osada, N., Hwang, C.-C., Chiang, T.-Y., 2012. Evolutionary rates of commonly used nuclear and organelle markers of *Arabidopsis relatives* (Brassicaceae). *Gene* 499 (1), 194–201. <https://doi.org/10.1016/j.gene.2012.02.037>.
- Hwang, Y., Kim, M., 2012. A taxonomic study of *Hemerocallis* (Liliaceae) in Korea. *Korean J. Pl. Taxon.* 42, 294–306. <https://doi.org/10.11110/kjpt.2012.42.4.294> (in Korean with English abstract).
- Ito, T., Nakanishi, H., Chichibu, Y., Minoda, K., Kokubugata, G., 2017. *Sedum danjoense* (Crassulaceae), a new species of succulent plants from the Danjo Islands in Japan. *Phytotaxa* 309, 23–34. <https://doi.org/10.11646/phytotaxa.309.1.2>.
- Kang, S.S., Chung, M.G., 1997a. Genetic variation and population structure in Korean endemic species: IV. *Hemerocallis hakuensis* (Liliaceae). *J. Plant Res.* 110 (2), 209–217. <https://doi.org/10.1007/BF02509309>.



- Kang, S.S., Chung, M.G., 1997b. *Hemerocallis taeanensis* (Liliaceae), a new species from Korea. *Syst. Bot.* 22, 427–431. <https://doi.org/10.2307/2419818>.
- Kato, M., Morita, S.-I., Yamaguchi, T., Kako, M., 1967. Note on the fauna of the Danjo islands, Nagasaki prefecture, Japan. *Reports Cultural Properties, Nagasaki Prefecture* 6, 59–97 (in Japanese).
- Kawano, S., 1961. On the natural hybrid population of *Hemerocallis*. *Can. J. Bot.* 39 (3), 667–681.
- Kawano, S., Noguchi, J., 1973. Biosystematic studies on the genus *Hemerocallis* (Liliaceae). I. Introgressive hybridization between *H. citrina* v. *vespertina* and *H. fulva* sensu lato. *J. Coll. Lib. Arts Toyama Univ.* 6, 111–167.
- Kay, K.M., Whittall, J.B., Hodges, S.A., 2006. A survey of nuclear ribosomal internal transcribed spacer substitution rates across angiosperms: an approximate molecular clock with life history effects. *BMC Evol. Biol.* 6, 36. <https://doi.org/10.1186/1471-2148-6-36>.
- Klahre, U., Gurba, A., Hermann, K., Saxenhofer, M., Bossolini, E., Guerin, P.M., Kuhlmeier, C., 2011. Pollinator choice in *Petunia* depends on two major genetic loci for floral scent production. *Curr. Biol.* 21, 730–739. <https://doi.org/10.1016/j.cub.2011.03.059>.
- Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A., Mayrose, I., 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* 15, 1179–1191. <https://doi.org/10.1111/1755-0998.12387>.
- Larson, W.A., Seeb, L.W., Everett, M.V., Waples, R.K., Templin, W.D., Seeb, J.E., 2014. Genotyping by sequencing resolves shallow population structure to inform conservation of Chinook salmon (*Oncorhynchus tshawytscha*). *Evol. Appl.* 7 (3), 355–369. <https://doi.org/10.1111/eva.12128>.
- Lassmann, T., Hayashizaki, Y., Daub, C.O., 2009. TagDust—a program to eliminate artifacts from next generation sequencing data. *Bioinformatics* 25 (21), 2839–2840. <https://doi.org/10.1093/bioinformatics/btp527>.
- Lelej, A.S., Kupianskaya, A.N., 2000. The bumble-bees (Hymenoptera, Apidae, Bombinae) of the Kuril Islands. *Far East. Entomol.* 95, 1–17.
- Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25 (14), 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25 (16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
- Li, L.F., Wang, H.Y., Pang, D., Liu, Y., Liu, B., Xiao, H.X., 2014. Phenotypic and genetic evidence for ecological speciation of *Aquilegia japonica* and *A. oxysepala*. *New Phytol.* 204, 1028–1040. <https://doi.org/10.1111/nph.12967>.
- Li, M.R., Wang, H.Y., Ding, N., Lu, T., Huang, Y.C., Xiao, H.X., Liu, B., Li, L.F., 2019. Rapid divergence followed by adaptation to contrasting ecological niches of two closely related columbine species *Aquilegia japonica* and *A. oxysepala*. *Genome Biol. Evol.* 11, 919–930. <https://doi.org/10.1093/gbe/evz038>.
- López-Pujol, J., Zhang, F.-M., Sun, H.-Q., Ying, T.-S., Ge, S., 2011. Centres of plant endemism in China: places for survival or for speciation? *J. Biogeogr.* 38, 1267–1280. <https://doi.org/10.1111/j.1365-2699.2011.02504.x>.
- Luu, K., Bazin, E., Blum, M.G., 2017. pcdapt: an R package to perform genome scans for selection based on principal component analysis. *Mol. Ecol. Resour.* 17, 67–77. <https://doi.org/10.1111/1755-0998.12592>.
- Martin, M., Patterson, M., Garg, S., Fischer, S.O., Pisanti, N., Klau, G.W., Schoenhuth, A., Marschall, T., 2016. WhatsHap: fast and accurate read-based phasing. *bioRxiv*, 085050. <https://doi.org/10.1101/085050>.
- Matsumoto, T., Yasumoto, A.A., Nitta, K., Hirota, S.K., Yahara, T., Tachida, H., 2015. Difference in flowering time can initiate speciation of nocturnally flowering species. *J. Evol. Biol.* 370, 61–71. <https://doi.org/10.1016/j.jtbi.2015.01.036>.
- Matsumoto, T., Yasumoto, A.A., Nitta, K., Yahara, T., Tachida, H., 2013. Difference in flowering time as an isolating barrier. *J. Theor. Biol.* 317, 161–167. <https://doi.org/10.1016/j.jtbi.2012.10.001>.
- Matsuoka, M., Hotta, M., 1966. Classification of *Hemerocallis* in Japan and its vicinity. *Acta Phytotax. Geobot.* 22, 25–43 (in Japanese with English key to the sections and species and English taxonomic treatments).
- McLay, T.G.B., Bayly, M.J., 2016. A new family placement for Australian blue squill, *Chamaescilla*: Xanthorrhoeaceae (Hemerocallidoideae), not Asparagaceae. *Phytotaxa* 275, 97–111. <https://doi.org/10.11646/phytotaxa.275.2.2>.
- Meirmans, P.G., Van Tienderen, P.H., 2004. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Mol. Ecol. Notes* 4 (4), 792–794. <https://doi.org/10.1111/j.1471-8286.2004.00770.x>.
- Miyata, A., 1973. Moth collection at night in Danjo Island. *Yadoriga* 72, 3–10 (in Japanese).
- Murakami, S., Takayama, K., Fuse, S., Hirota, S.K., Koi, S., Ideno, T., Yamamoto, T., Tamura, M.N., 2020. Recircumscription of Sections of the Genus *Hemerocallis* (Asphodelaceae) from Japan and its adjacent regions based on MIG-seq data. *Acta Phytotax. Geobot.* 71, 1–11. <https://doi.org/10.18942/phyt.201913>.
- Nielsen, R., Wakeley, J., 2001. Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics* 158, 885–896.
- Nitta, K., Yasumoto, A.A., Yahara, T., 2010. Variation of flower opening and closing times in F1 and F2 hybrids of daylily (*Hemerocallis fulva*; Hemerocallidaceae) and nightlily (*H. citrina*). *Am. J. Bot.* 97 (2), 261–267. <https://doi.org/10.3732/ajb.0900001>.
- Noguchi, J., De-yuan, H., 2004. Multiple origins of the Japanese nocturnal *Hemerocallis citrina* var. *vespertina* (Asparagales: Hemerocallidaceae): Evidence from noncoding chloroplast DNA sequences and morphology. *Int. J. Plant Sci.* 165 (1), 219–230. <https://doi.org/10.1086/380983>.
- Noguchi, J., Hong, D.Y., Grant, W.F., 2004. The historical evolutionary development of *Hemerocallis middendorffii* (Hemerocallidaceae) revealed by non-coding regions in chloroplast DNA. *Plant Syst. Evol.* 247, 1–22. <https://doi.org/10.1007/s00606-004-0139-x>.
- Ou, X., Liu, G.e., Wu, L.-H., 2020. The complete chloroplast genome of *Hemerocallis citrina* (Asphodelaceae), an ornamental and medicinal plant. *Mitochondrial DNA Part B* 5 (1), 1109–1110. <https://doi.org/10.1080/23802359.2020.1726227>.
- Owen, C.R., Bradshaw, H.D., 2011. Induced mutations affecting pollinator choice in *Mimulus lewisii* (Phrymaceae). *Arthropod-Plant Interactions* 5 (3), 235–244. <https://doi.org/10.1007/s11829-011-9133-8>.
- Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
- Quattrocchio, F., Wing, J., van der Woude, K., Souer, E., de Vetten, N., Mol, J., Koes, R., 1999. Molecular analysis of the *anthocyanin2* gene of *Petunia* and its role in the evolution of flower color. *Plant Cell* 11 (8), 1433–1444.
- Rieseberg, L.H., 1997. Hybrid origins of plant species. *Annu. Rev. Ecol. Syst.* 28 (1), 359–389. <https://doi.org/10.1146/annurev.ecolsys.28.1.359>.
- Rieseberg, L.H., Soltis, D.E., 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. *Evol. Trends Plants* 5, 65–84.
- Rieseberg, L.H., Van Fossen, C., Desrochers, A.M., 1995. Hybrid speciation accompanied by genomic reorganization in wild sunflowers. *Nature* 375 (6529), 313–316. <https://doi.org/10.1038/375313a0>.
- Rodriguez-Enriquez, M.J., Grant-Downton, R.T., 2013. A new day dawning: *Hemerocallis* (daylily) as a future model organism. *AoB Plants* 5, pls055. <https://doi.org/10.1093/aobpla/pls055>.
- Roesti, M., Salzburger, W., Berner, D., 2012. Uninformative polymorphisms bias genome scans for signatures of selection. *BMC Evol. Biol.* 12 (1), 94. <https://doi.org/10.1186/1471-2148-12-94>.
- Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E., Sanchez-Gracia, A., 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* 34, 3299–3302. <https://doi.org/10.1093/molbev/msx248>.
- Sagawa, J.M., Stanley, L.E., LaFountain, A.M., Frank, H.A., Liu, C., Yuan, Y.W., 2016. An R2R3-MYB transcription factor regulates carotenoid pigmentation in *Mimulus lewisii* flowers. *New Phytol.* 209, 1049–1057. <https://doi.org/10.1111/nph.13647>.
- Schiestl, F.P., Johnson, S.D., 2013. Pollinator-mediated evolution of floral signals. *Trends Ecol. Evol.* 28 (5), 307–315. <https://doi.org/10.1016/j.tree.2013.01.019>.
- Smit, A.F.A., Hubley, R., Green, P., 2015. RepeatMasker. <http://www.repeatmasker.org> (accessed 19 May 2019).
- Spaethe, J., Tautz, J., Chittka, L., 2001. Visual constraints in foraging bumblebees: flower size and color affect search time and flight behavior. *Proc. Natl. Acad. Sci. USA* 98 (7), 3898–3903. <https://doi.org/10.1073/pnas.071053098>.
- Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>.
- Strasburg, J.L., Rieseberg, L.H., 2008. Molecular demographic history of the annual sunflowers *Helianthus annuus* and *H. petiolaris*—large effective population sizes and rates of long-term gene flow. *Evolution* 62, 1936–1950. <https://doi.org/10.1111/j.1558-5646.2008.00415.x>.
- Strasburg, J.L., Rieseberg, L.H., 2010. How robust are “isolation with migration” analyses to violations of the im model? A simulation study. *Mol. Biol. Evol.* 27 (2), 297–310. <https://doi.org/10.1093/molbev/msp233>.
- Streisfeld, M.A., Rausher, M.D., 2009. Genetic changes contributing to the parallel evolution of red floral pigmentation among *Ipomoea* species. *New Phytol.* 183 (3), 751–763. <https://doi.org/10.1111/j.1469-8137.2009.02929.x>.
- Suarez-Gonzalez, A., Lexer, C., Cronk, Q.C.B., 2018. Adaptive introgression: a plant perspective. *Biol. Lett.* 14 (3), 20170688. <https://doi.org/10.1098/rsbl.2017.0688>.
- Sugiura, N., 1978. Further analysts of the data by akaike’s information criterion and the finite corrections. *Commun. Stat. - Theory Methods* 7 (1), 13–26. <https://doi.org/10.1080/03610927808827599>.
- Suyama, Y., Matsuki, Y., 2015. MIG-seq: an effective PCR-based method for genome-wide single-nucleotide polymorphism genotyping using the next-generation sequencing platform. *Sci. Rep.* 5, 16963. <https://doi.org/10.1038/srep16963>.
- Takahashi, K., Terai, Y., Nishida, M., Okada, N., 2001. Phylogenetic relationships and ancient incomplete lineage sorting among cichlid fishes in Lake Tanganyika as revealed by analysis of the insertion of retrotransposons. *Mol. Biol. Evol.* 18, 2057–2066. <https://doi.org/10.1093/oxfordjournals.molbev.a003747>.
- Tanabe, A.S., 2007. kakusan: a computer program to automate the selection of a nucleotide substitution model and the configuration of a mixed model on multilocus data. *Mol. Ecol. Notes* 7 (6), 962–964. <https://doi.org/10.1111/j.1471-8286.2007.01807.x>.
- Toriba, M., 1986. Preliminary study on the systematic status of a Danjo Islands snake. *Japanese J. Herpetol.* 11 (3), 124–136. <https://doi.org/10.5358/hsj1972.11.3.124>.
- Uematsu, S., Iwamoto, S., Syokyu, N., Nakamura, R., Nishimura, T., Hayashida, M., Hirano, Y., Motoike, H., Yoshitomi, K., 1973. Plants of Danjo Islands. In: Society, N. B. (Ed.), *Fauna and Flora of the Danjo Islands*. Nagasaki Biological Society, Nagasaki, pp. 25–73 (in Japanese).
- van der Niet, T., Johnson, S.D., 2012. Phylogenetic evidence for pollinator-driven diversification of angiosperms. *Trends Ecol. Evol.* 27 (6), 353–361. <https://doi.org/10.1016/j.tree.2012.02.002>.
- van der Pijl, L., 1961. Ecological aspects of flower evolution. II. Zoophilous flower classes. *Evolution* 15 (1), 44–59.
- Wagner, C.E., Keller, I., Wittwer, S., Selz, O.M., Mwaiko, S., Greuter, L., Sivasundar, A., Seehausen, O., 2013. Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid



- adaptive radiation. *Mol. Ecol.* 22 (3), 787–798. <https://doi.org/10.1111/mec.12023>.
- Whittall, J.B., Hodges, S.A., 2007. Pollinator shifts drive increasingly long nectar spurs in columbine flowers. *Nature* 447 (7145), 706–709.
- Wolf, P.G., Campbell, D.R., Waser, N.M., Sipes, S.D., Toler, T.R., Archibald, J.K., 2001. Tests of pre- and postpollination barriers to hybridization between sympatric species of *Ipomopsis* (Polemoniaceae). *Am. J. Bot.* 88 (2), 213–219. <https://doi.org/10.2307/2657012>.
- Yamaguchi, T., Ejima, M., 1973. Overview of Danjo Islands. In: Society, N.B. (Ed.), *Fauna and Flora of the Danjo Islands*. Nagasaki Biological Society, Nagasaki, pp. 1–21 (in Japanese).
- Yamaguchi, T., Miyagi, I., Miyata, A., Noda, M., Eshima, M., Yoshida, K., 1973. List of insects of the Danjo islands. In: Society, N.B. (Ed.), *Fauna and Flora of the Danjo Islands*. Nagasaki Biological Society, Nagasaki, pp. 88–101 (in Japanese).
- Yasumoto, A.A., Yahara, T., 2006. Post-pollination reproductive isolation between diurnally and nocturnally flowering daylilies, *Hemerocallis fulva* and *Hemerocallis citrina*. *J. Plant Res.* 119 (6), 617–623. <https://doi.org/10.1007/s10265-006-0028-1>.
- Yasumoto, A.A., Yahara, T., 2008. Reproductive isolation on interspecific backcross of F1 pollen to parental species, *Hemerocallis fulva* and *H. citrina* (Hemerocallidaceae). *J. Plant Res.* 121 (3), 287–291. <https://doi.org/10.1007/s10265-008-0152-1>.